NORTH-HOLLAND
MATHEMATICS STUDIES 159
Editor: Leopoldo NACHBIN

Approximation Problems in Analysis and Probability

M.P. HEBLE

NORTH-HOLLAND

APPROXIMATION PROBLEMS IN ANALYSIS AND PROBABILITY

NORTH-HOLLAND MATHEMATICS STUDIES 159

(Continuation of the Notas de Matemática)

Editor: Leopoldo NACHBIN

Centro Brasileiro de Pesquisas Fisicas
Rio de Janeiro, Brazil
and
University of Rochester
New York, U.S.A.

NORTH-HOLLAND – AMSTERDAM • NEW YORK • OXFORD • TOKYO

APPROXIMATION PROBLEMS IN ANALYSIS AND PROBABILITY

M.P. HEBLE

Department of Mathematics
University of Toronto
Toronto, Canada

1989

NORTH-HOLLAND – AMSTERDAM • NEW YORK • OXFORD • TOKYO

ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the U.S.A. and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
655 Avenue of the Americas
New York, N.Y. 10010, U.S.A.

Library of Congress Cataloging-in-Publication Data

Heble, M. P.
Approximation problems in analysis and probability / M.P. Heble.
p. cm. -- (North-Holland mathematics studies ; 159)
Includes bibliographical references.
ISBN 0-444-88021-6
1. Approximation theory. 2. Mathematical analysis.
3. Probabilities. I. Title. II. Series.
QA221.H375 1989
511'.4--dc20 89-16147
CIP

ISBN: 0 444 88021 6

Printed in the Netherlands

To

My mother Girijabai
My uncle Rama Rao
and
Sushila, Ajay and Sucheta

This Page Intentionally Left Blank

Table of Contents

This Page Intentionally Left Blank

Introduction

The classical Weierstrass-Stone theorem and the Bernstein-type weighted approximation theorems were greatly extended by L. Nachbin. Another aspect of approximation theory, now called strong approximation and initiated by H. Whitney, had simultaneously developed, with contributions in finite-dimensional spaces as also in infinite-dimensional spaces, by various individuals. At the same time, several approximation results in a probabilistic setting - from the elegant probabilistic proof of Weierstrass' theorem by S. Bernstein to the later results on convergence of stochastic processes established by A.V. Skorokhod and other later authors - were being added to the literature.

The material in this book covers some special aspects of the approximation theory of functions, viz. strong approximation in function spaces, as also certain approximation results concerning stochastic processes. The choice of topics reflects only the author's taste. Within the narrow range of topics chosen, I have tried to do as thorough justice as I could, to the subject as also to the contribution of various individuals active in their respective areas; any possible omission of names is unintentional.

This book is meant to be a monograph, of interest to research workers in the fields of analysis, probability, and stochastic processes. Graduate students, hopefully, will find it useful not merely as a source of information but also as an incentive to spur then on to do further work. The author has noted other monographs recently published, covering related topics. However, the contents of these books show that the overlap between these and my present monograph is negligible (e.g., cf. K. Sundaresan and S. Swami-

nathan: "Geometry and non-linear analysis in Banach spaces", Springer Verlag Lecture Notes in Math. No. 1131, 1986; and J.G. Llavona: "Approximation of continuously differentiable functions", Notas de Matematica No. 130, North Holland 1986).

A quick description of the contents of this book appears to be in order. The material is divided into four chapters. The first chapter gives a quick survey of the classical Weierstrass-Stone theorem, Bernstein's weighted approximation problem and Nachbin's extension of the classical Bernstein approximation results. The material in this chapter excluding sections 4 and 5, is mostly a summary of Professor Leopoldo Nachbin's monograph: "Elements of approximation theory". In Chapter II we present strong approximation results in a finite-dimensional space $\mathbf{R}^n$ - first H. Whitney's theorem on strong approximation by real analytic functions, and then some results on C^∞-approximation (strong sense); the latter appear to have been commonly known and there are excellent expositions in several monographs, hence we have been content with only a summary in this book. Chapters III presents strong approximation results in finite- or infinite-dimensional separable spaces, starting with Kurzweil's extension of Whitney's theorem (on analytic approximation), and ending with some recent results established by this author, as also an indication of possibilities in other directions. We also explain a connection between strong approximation results and the earlier Bernstein-Nachbin ideas. In Chapter IV we present some probabilistic approximation results, starting with a quick look at Bernstsein's well-known proof of Weierstrass' theorem with some recent developments, and ending with some results by A.V. Skorokhod on approximation of stochastic processes. Here, again, individual choice was the guiding factor. We thought it necessary

to leave out the enormous area of weak approximation - an area which has found excellent exposition in several monographs, e.g., M. Rosenblatt: "Markov processes, structure and asymptotic properties" (Springer Verlag 1971), and D. Pollard: "Convergence of stochastic processes", (Springer Verlag 1984).

As for organisation of the the material, theorems, lemmas, etc., are numbered according to chapter and section; thus Theorem II 2.1 means Theorem 1 in section 2 of Chapter II. Equations and formulae are numbered consecutively, but the numbering is separate for each section and each chapter. We have used the common symbols: "$\exists$" for "there exists", "$\forall$" for "for all" or "for any", "$\ni$" for "such that", "$\Rightarrow$" for "implies", and "$\Leftrightarrow$" for "if and only if", $\mathbf{C}$ denotes the set of complex numbers, $\mathbf{R}$ the set of real numbers and $\mathbf{R}^n$ the n-dimensional real Euclidean space. There are four appendices at the end of the book explaining basic background material without proofs, and with sufficient further references.

The writing of this monograph was partially supported by an NSERC operating grant. Thanks are due to Shirley Chan and Pat Broughton for patient and expert typing of the manuscript. I am indebted to Professor Leopoldo Nachbin, first, for encouraging me to write this book for the Notas de Matematica series, and secondly for permission to summarise the material of his monograph: "Elements of approximation theory".

This Page Intentionally Left Blank

CHAPTER I

Weierstrass-Stone theorem and generalisations - a brief survey

In the first three sections of this chapter we shall review known results concerning the classical Weierstrass-Stone theorem (cf. [60]), Bernstein's approximation problem, (cf. [41]) and further generalisations by L. Nachbin. The material in this chapter, excluding sections 4 and 5, is taken from L. Nachbin's published lecture notes [39] and for this reason we shall often present only a summary, leaving it to the reader to refer to his monograph for further details. Any missing proofs will be found in [43]. For convenience in presentation, the chapter is divided into five subsections. For the concepts of functional analysis used we refer the reader to Appendix 1.

Throughout this chapter we shall assume E to be a completely regular space, i.e., a Hausdorff topological space such that for any $a \in E$, and any closed subset $F \subset E$ not containing a, there is a continuous real-valued function f on E such that $0 \leq f \leq 1$, $f(a) = 1$ and $f(x) = 0$ for any $x \in F$.

We shall denote by $C(E; \mathbb{C})$ the commutative algebra with unit of all continuous complex-valued functions on E; if $E = \phi$ we set $C(E; \mathbb{C}) = \{0\}$. For convenience we shall often write $C(E)$ for $C(E; \mathbb{C})$. Every compact set $K \subset E$ determines an algebra semi-norm $\|f\|_K \stackrel{\text{def}}{=} \max\{|f(x)|, x \in K\}$ on $C(E)$. Let Γ be the family of such semi-norms (cf. Appendix 1). We shall understand that $C(E)$ is endowed with the *compact-open topology* viz. the topology τ_Γ determined by the family Γ of semi-norms. Then $C(E)$ becomes a topological algebra, i.e., a topological vector-space which is also an algebra. If E is compact then $C(E)$ becomes a Banach algebra.

We note that the topology τ_r is the topology of uniform convergence on compact sets. This fact will be utilised in certain proofs. In proving results for a $C(E)$ with E completely regular, it will be often convenient to prove any one such result on the assumption that E is a compact space, the result will then follow for a completely regular space. Sometimes it may be necessary to use the result: every element in $C(K)$, where $K \subset E$ is compact, has an extension to $C(E)$.

The subalgebra of $C(E)$ consisting of real-valued continuous functions on E will be denoted by $C(E; \mathbf{R})$, and will be endowed with the subspace topology inherited from τ_r.

§1. Weierstrass-Stone theorem.

The first non-trivial theorem that we should note is the following result, due to S. Kakutani and M.H. Stone. A subset $L \subset C(E; \mathbf{R})$ is called a *lattice* if: $f, g \in L \Rightarrow \sup(f, g) \in L$ and $\inf(f, g) \in L$.

Theorem 1.1. *(Kakutani-Stone) Let $L \subset C(E; \mathbf{R})$ be a lattice and $f \in C(E; \mathbf{R})$. Then f belongs to the closure of $L \Leftrightarrow \forall x_1, x_2 \in E$, and $\forall \varepsilon > 0, \exists g \in L \ni$*

$$|g(x_1) - f(x_1)| < \varepsilon, \qquad |g(x_2) - f(x_2)| < \varepsilon \ . \tag{1}$$

Proof. Only the sufficiency part of the assertion meeds some attention. As noted earlier, it is enough to prove the statement on the assumption that E is compact. So suppose E is compact. Let $f \in C(E; \mathbf{R})$ satisfy the condition (1), and let $\varepsilon > 0$. Then, using compactness, it follows that for any given $t \in E, \exists g_t \in L$ satisfying:

$$\begin{cases} g_t(u) > f(u) - \varepsilon & \text{for any} \quad u \in E, \quad \text{and} \\ g_t(t) < f(t) + \varepsilon \ . \end{cases}$$

Then again using compactness, we find that $\exists g \in L$ satisfying:

$$f(u) - \varepsilon < g(u) < f(u) + \varepsilon \qquad \text{for any } u \in E\ .$$

Thus $\|g - f\|_E < \varepsilon$, and hence $g \in \bar{L}$. This completes the proof.

The next step is to note a result concerning the closure of an ideal $I \subset C(E)$. An *ideal* in $C(E)$ is a nonempty subset $I \subset C(E)$ such that for any $g \in I$, $fg \in I \quad \forall f \in C(E)$.

Proposition 1.2. *Let $I \subset C(E)$ be an ideal, and let $f \in C(E)$. Suppose N is the closed subset of E consisting of all $x \in E \ni g(x) = 0\ \forall g \in I$. Then $f \in \bar{I} \Leftrightarrow f = 0$ on N.*

Proof. We note that $N = \cap_{f \in I} f^{-1}(\{0\})$, hence N is closed. Also for each $x \in E$ the delta function δ_x is continuous (cf. Appendix). For each $x \in N, \delta_x$ vanishes on I, hence $\delta_x = 0$ on $\bar{I}$. This proves "$\Rightarrow$".

Next suppose E is compact and suppose $f \in C(E)$ vanishes on N. Let $\varepsilon > 0$, and define $X = \{x \in E \mid |f(x)| \geq \varepsilon\}$. Then X is closed, and $X \cap N = \phi$. Now let $h \in I \ni 0 \leq h \leq 1$, and $h(x) = 1 \forall x \in X$ (one such h does exist). Now set $g = fh \in I$. Then $\|f - g\|_E < \varepsilon$, and hence $f \in \bar{I}$. This proves "$\Leftarrow$". The proposition follows.

We turn now to the Weierstrass-Stone theorem. A subset $X \subset C(E; \mathbb{C})$ is said to be *self-adjoint* if $f \in X \Rightarrow \bar{f} \in X$, where $\bar{f}$ is the complex conjugate of f.

Theorem 1.3. (Weierstrass-Stone) *Suppose $A \subset C(E)$ is a subalgebra, which we assume to be self-adjoint in the complex case. Let $f \in C(E)$. Then $f \in \bar{A} \Leftrightarrow$*

(1) $\forall x_1, x_2 \in E \ni \ f(x_1) \neq f(x_2), \exists g \in A \ni g(x_1) \neq g(x_2)$;

(2) $\forall x \in E \ni \ f(x) \neq 0,\ \exists g \in A \ni g(x) \neq 0$.

For the proof the following lemmas are needed.

Lemma 1.4. *If $k \geq 0$, and $\varepsilon > 0$, then $\exists$ polynomial $p : \mathbf{R} \to \mathbf{R}, \ni p(0) = 0$ and $| \, |p(t) - |t| \, |< \varepsilon \forall t \ni |t| \leq k$.*

Proof. Since the statement is trivial for $k = 0$, we shall suppose $k > 0$. Also if the Lemma is true for $k = 0$ then it is true for any $k > 0$. For suppose $| \; p(t) - |t| \; |< \varepsilon$ for $|t| \leq 1$; then $\left|kp(\frac{t}{k}) - |t|\right| < k\varepsilon$ for $|t| \leq k$. Hence we shall assume $k = 1$. Now we define the polynomials p_n, for $n = 0, 1, 2, \dots$, on $\mathbf{R}$, by:

$$p_0 \;=\; 0, p_{n+1}(t) = p_n(t) + \frac{1}{2}[t^2 - p_n(t)^2], \quad t \in \mathbf{R} \;.$$

Then $p_n(0) = 0$, and by induction it follows that

$$0 \leq p_n(t) \leq |t| \quad \text{if} \quad |t \leq 1 \quad \text{for} \quad n = 0, 1, 2, \dots \;. \tag{2}$$

Next for $n \geq 1$ we find (by induction)

$$|t| - p_{n+1}(t) \leq [|t| - p_n(t)](1 - \frac{|t|}{2} \quad \text{for} \quad |t| \leq 1 \;,$$

and hence for $n = 0, 1, 2, \dots,$

$$|t| - p_n(t) \leq |t|\big(1 - \frac{|t|}{2}\big)^n \quad \text{for} \quad |t| \leq 1 \;. \tag{3}$$

Now let $\varepsilon \;\ni\; 0 < \varepsilon \leq 1$. Then from (2) and (3) it follows that $\exists$ integer $n_0 \geq 0 \ni$

$$0 \leq |t| - p_n(t) \leq |t| \leq \varepsilon \quad \text{if} \quad |t| \leq \varepsilon \;,$$

$$0 \leq |t| - p_n(t) \leq \big(1 - \frac{\varepsilon}{2}\big)^n \leq \varepsilon \quad \text{if} \quad \varepsilon \leq |t| \leq 1, n \geq n_0 \;.$$

Hence $0 \leq |t| - p_n(t) \leq \varepsilon$ if $|t| \leq 1, n \geq n_0$. Thus the lemma is proved.

Lemma 1.5. **(Lebesgue)** *Every closed subalgebra $A \subset C(E; \mathbf{R})$ is a lattice.*

Lemma 1.6. *Let $A \subset \mathbf{R}^2$ be a subalgebra and $b \in \mathbf{R}^2$. Then $b \notin A \Leftrightarrow$ at least one of the following conditions holds:*

(1) $b \notin \mathbf{R} \times 0$ *and* $A \subset \mathbf{R} \times 0$;

(2) $b \in 0 \times \mathbf{R}$ *and* $A \subset 0 \times \mathbf{R}$;

(3) $b \notin \Delta$ *and* $A \subset \Delta$.

Proof of Theorem 1.3. Again we shall attend only to the sufficiency part of the proof. Suppose the conditions (1) and (2) of the theorem hold. If $x_1, x_2 \in E$ the mapping $g \in C(E;\mathbf{R}) \to \big(g(x_1), g(x_2)\big) \in \mathbf{R}^2$ is an algebra homomorphism, which we shall call Φ. By (2) if $\Phi(f) \notin \mathbf{R} \times O$, then $\Phi(A) \not\subset 0 \times \mathbf{R}$. Again by (1), if $\Phi(f) \notin \Delta$, then $\Phi(A) \not\subset \Delta$. Then by Lemma 1.6, $\Phi(f) \in \Phi(A)$ i.e., $\exists g \in A \subset \bar{A}$; $\bar{A}$ is a lattice by Lemma 1.5, hence $f \in \overline{\overline{A}}$ (using Theorem 1.1), i.e., $f \in \bar{A}$. Thus the theorem is proved.

Next, using the result of the next lemma, Theorem 1.3 follows in the complex case as well.

Lemma 1.7. *Let $A \subset C(E;\mathbf{C})$ be a self-adjoint algebra. Then the set ReA of the real parts Ref of all $f \in A$ is a subalgebra of $C(E;\mathbf{R})$ and $A = ReA + iReA$.*

The following is a consequence of Theorem 1.3.

Corollary 1.8. *Suppose $A \subset C(E)$ is a subalgebra which we assume to be self-adjoint in the complex case. Then A is dense in $C(E) \Leftrightarrow$*

(1) A is "separating on E" i.e., $\forall x_1, x_2 \in E$ with $x_1 \neq x_2$, $\exists g \in A \ni g(x_1) \neq g(x_2)$;

(2) A is "non-vanishing on E", i.e., $\forall x \in E, \exists g \in A \ni g(x) \neq 0$.

We also note here that the classical Weierstrass theorem follows as a consequence of the Weierstrass-Stone theorem.

§2. Closure of a module, the weighted approximation problem.

The closure theorems of the preceding section can be seen to be special instances of closure theorems concerning modules; this is the topic of this section.

Given a set A of C-valued functions on a set E, we introduce an equivalence relation on E, denoted E/A, as follows: if $x_1, x_2 \in E$, the $x_1 \sim x_2$ modulo E/A if $f(x_1) = f(x_2) \forall f \in A$.

Theorem 2.1. *Suppose $A \subset C(E)$ is a subalgebra containing the unit, and is assumed to be self-adjoint in the complex case, $W \subset C(E)$ is a vector subspace which is also a module over A, i.e., $AW \subset W$; and let $f \in C(E)$. Then $f \in \overline{W} \in C(E) \Leftrightarrow \forall$ equivalence class $X \subset E$ modulo $E/A, \forall$ compact set $K \subset X$, and $\forall \varepsilon > 0 \exists w \in W \ni |w(x) - f(x)| < \varepsilon \forall x \in K$.*

The proof requires the existence of a "continuous partition of unity", and the following result is just enough for this purpose at the moment; later on we shall note a stronger result on partitions of unity. A *continuous partition of unity subordinate to a finite open covering* $V_1 \cup \ldots \cup V_n$ of E is a finite sequence $f_1, \ldots, f_n \in C(E; \mathbf{R}) \ni$ each $f_i \geq 0$, and is 0 outside V_i, and $\sum_{i=1}^{n} f_i = 1$.

Lemma 2.2. (Dieudonné and Bochner) *Suppose E is a normal space. Then $\exists$ continuous partition of unity on E subordinated to any finite open covering of E.*

Proof of Theorem 2.1. If $f \in \overline{W}$ then clearly f satisfies the conditions stated in the theorem. Conversely suppose the condition stated in the theorem holds. We shall assume E compact. Let F be the quotient space of E modulo the equivalence relation $E/A, \pi$ the natural projection $E \to F$, and we shall understand that F is endowed with

the quotient topology. For every $f \in A$ define g on F by: $g\pi = f$. The mapping $f \to g$ defines a mapping $A \to C(F)$; let B be the image of A under this mapping. Then B is separating on F, which we note to be a compact Hausdorff space.

Now let $\varepsilon > 0$, and $y \in F$. Then $\pi^1(\{y\}) \subset E$ is an equivalence class and is compact. By assumption $\exists w_y \in W \ni |w_y(x) - f(x)| < \varepsilon$ for $x \in \pi^{-1}(\{y\})$. The set $K_y = \{x \in E \mid |w_y(x) - f(x)| \geq \varepsilon\} \subset E$ is compact and $\pi^{-1}(\{y\}) \cap K_y = \emptyset$, hence $y \notin \pi(K_y)$. It follows that $\bigcap_{y \in F} \pi(K_y) = \emptyset$, and by the finite intersection property $\exists y_1, \ldots, y_n \in F \ni \pi(K_{y_1}) \cap \ldots \cap \pi(K_{y_n}) = \emptyset$. By Lemma 2.2 $\exists \psi_1, \ldots, \psi_n \in C(F) \ni$ each $\psi_i \geq 0$ and $= 0$ on $\pi(K_{y_i})$, $i = 1, \ldots, n$, and $\psi_1 + \cdots + \psi_n = 1$. Then $\phi_i = \psi_i \pi$, $i = 1, \ldots, n$, satisfy: $\phi_i \geq 0$ and $= 0$ on K_{y_i}, $i = 1, \ldots, n$, and $\phi_1 + \cdots + \phi_n = 1$.

Now a little argument shows that

$$\left|\sum_i \phi_i(x) w_{y_i}(x) - f(x)\right| \leq \varepsilon \qquad \forall x \in E\ . \tag{1}$$

We note that each ϕ_i is constant on every equivalence class modulo E/A; the subalgebra A contains 1, and is self-adjoint in the complex case; thus $\phi_i \in \bar{A}$ in $C(E)$. Hence $\forall \delta > 0$ $\exists h_i \in A$ $(i = 1, \ldots, n) \ni$

$$|h_i(x) - \phi_i(x)| \leq \delta \qquad \forall x \in E, i = 1, \ldots, n\ . \tag{2}$$

Let $M = \sup\{|w_{y_i}(x)| \mid x \in E, i = 1, \ldots, n\}$, and choose $\delta \leq \frac{\varepsilon}{Mn}$. Then from (4) and (5) we find

$$|\sum h_i(x) w_{y_i}(x) - f(x)| \leq 2\varepsilon \qquad \forall x \in E\ .$$

Since $h_i w_{y_i} \in AW \subset W, i = 1, \ldots, n$, therefore $f \in \overline{W}$. This completes the proof.

We shall simply state the theorems of Dieudonné and Choquét-Deny on closure in tensor products and on the closure of a convex sup-lattice, respectively.

Theorem 2.3. (Dieudonné) *Suppose E and F are completely regular spaces, and let $f \in C(E \times F)$. Then $\forall$ compact $K \subset E \times F$, and $\forall \varepsilon > 0$, $\exists g_1, \ldots, g_m \in C(E)$, and $\exists h_1, \ldots, h_m \in C(F) \ni$*

$$\left| \sum_{i=1}^{m} g_i(x) h_i(y) - f(x,y) \right| \leq \varepsilon \qquad \forall (x,y) \in K .$$

Before stating the theorem of Choquét-Deny, some definitions are in order. A subset $S \subset C(E; \mathbf{R})$ is called a *sup-lattice* if $f, g \in S \Rightarrow \sup\,(f,g) \in S$, and is called an *inf-lattice* if $f, g \in S \Rightarrow \inf(f,g) \in S$. If ϕ, ψ are continuous linear functionals on $C(E; \mathbf{R})$, we write $\phi \geq \psi$ to mean $\phi(f) \geq \psi(f) \forall f \in C(E; \mathbf{R}) \ni f \geq 0$.

If $\phi \geq 0$, we say that ϕ is *positive.*

Theorem 2.4. (Choquét-Deny) *Suppose $S \subset C(E; \mathbf{R})$ is a sup-lattice and let $f \in C(E; \mathbf{R})$. Then $f \in \overline{S}$ (in $(C(E; \mathbf{R}))$ $\Leftrightarrow$*

(1) $\forall$ positive $\phi \in C(E; \mathbf{R})^*$ (the dual of $C(E; \mathbf{R})$ cf. Appendix 1) and $\forall a \in E$

$$\phi(f) - f(a) \geq \inf\bigl\{\phi(g) - g(a) \mid g \in S\bigr\} ;$$

(2) $\forall$ positive $\phi \in C(E; \mathbf{R})^*$, $\phi(f) \geq \inf\{\phi(g) \mid g \in S\}$.

At this point it is necessary to explain Bernstein's weighted approximation problem and for this purpose we should first explain the concept of a weighted locally convex space of continuous functions. The next step thereafter is to explain L. Nachbin's contribution towards extending the classical Bernstein approximation problem, viz. his work on the weighted approximation problem for modules.

We first turn to the concept of a weighted locally convex space. Let V be a set of upper-semicontinuous positive functions on E. We shall assume that V is *directed*, i.e., if

$v_1, v_2 \in V$, then $\exists\lambda > 0$ and $\exists v \in V \ni v_1 \leq \lambda v$ and $v_2 \leq \lambda v$. The elements of V are called *weights*. The vector subspace of $C(E)$ consisting of all $f \ni vf$ is bounded on E, for each $v \in V$, will be denoted by $CV_b(E)$. Then each $v \in V$ determines a semi-norm $p_v(f) = \sup\{v(x) \cdot |f(x)| \mid x \in E\}$ on $CV_b(E)$. We shall understand that $CV_b(E)$ is endowed with the natural topology i.e., the locally convex topology determined by the family of semi-norms $\{p_v(\cdot)\}_{v \in V}$. The vector subspace of $C(E)$ consisting of all $f \ni \forall v \in V$ and $\forall\varepsilon > 0$ the closed subset $\{x \in E \mid v(x) \cdot |f(x)| \geq \varepsilon\}$ is compact, will be denoted by $CV_\infty(E)$. It is clear that $CV_\infty(E) \subset CV_b(E)$, and the natural topology on $CV_\infty(E)$ is understood to be the topology induced by $CV_b(E)$.

A few observations are in order at this point. The family of semi-norms $\{p_v(\cdot)\}_{v \in V}$ in the preceding paragraph is directed because V itself is directed. If V consists of a single function $v(\cdot)$ then we shall denote $CV_b(E)$ and $CV_\infty(E)$ by $C_{v_b}(E)$ and $C_{v_\infty}(E)$, respectively, and if V consists of the constant function 1 then $C_{v_b}(E), C_{v_\infty}(E)$ will be denoted by $C_b(E)$, and $C_\infty(E)$ respectively. $CV_b(E)$ and $CV_\infty(E)$ are Hausdorff spaces if $\forall x \in E \exists v \in V \ni v(x) > 0$. $CV_b(E)$ is a module over $C_b(E)$, and $CV_\infty(E)$ is a sub module over $C_b(E)$. Furthermore, if $f \in CV_b(E)$, $g \in C(E)$, and $|g| \leq |f|$ then $g \in CV_b(E)$; a similar remark holds for $CV_\infty(E)$.

We further note the following: (i) if V is the set of characteristic functions of all compact subsets of E, then $C(E) = CV_b(E) = CV_\infty(E)$ as locally convex spaces; (ii) if V consists of just the constant function 1, then $CV_b(E) = C_b(E)$, and the topology on $C_b(E)$ is defined by the single norm $\|f\|_E$; also in this case $CV_\infty(E) = C_\infty(E)$, with the same norm $\|f\|_E$; (iii) if $E = \mathbf{R}^n$, and V consists of the $|p|$ for all C-valued polynomials

p on $\mathbf{R}^n$ then $CV_b(\mathbf{R}^n) = CV_\infty(\mathbf{R}^n) \neq C(\mathbf{R}^n)$; in this case if a norm $\|x\|$ is fixed on $\mathbf{R}^n$, and W consists of functions $(1 + \|x\|)^m$ for $m = 0, 1, 2, \ldots$, on $\mathbf{R}^n$ then it is known that $CW_b(\mathbf{R}^n) = CW_\infty(\mathbf{R}^n) = CV_b(\mathbf{R}^n) = CV_\infty(\mathbf{R}^n)$ as locally convex spaces; in this case the elements of $C(E)$ thus obtained are said to be *rapidly decreasing at infinity.*

We note the following result.

Proposition 2.5. *$C_b(E) \cap CV_\infty(E)$ is dense in $CV_\infty(E)$.*

For stating the next theorem of Dieudonné on dense subsets in tensor products some terminology should be explained. For each $i = 1, \ldots, n$ let E_i be a completely regular space, V_i a directed set of upper-semicontinuous positive functions on E_i; let $E = E_1 \times \cdots \times E_n$, and $V = V_1 \times V_2 \times \cdots \times V_n$. The following theorem of Dieudonné holds.

Theorem 2.6. **(Dieudonné)** *The set of all finite sums of tensor-products $f = f_1 \times f_2 \times \cdots \times f_n$, $f_i \in (CV_i)_\infty(E_i), i = 1, \ldots, n$, is dense in $CV_\infty(E)$.*

We turn now to the *Bernstein approximation problem* (cf. [43]). Let f be a C-valued function on $\mathbf{R}^n$, and suppose f is locally bounded, i.e., bounded on every compact subset of $\mathbf{R}^n$. Then f is said to be *rapidly decreasing at infinity* if the following equivalent conditions hold: (1) pf is bounded on $\mathbf{R}^n$ for any $p \in \mathcal{P}(\mathbf{R}^n)$ (the set of polynomials on $\mathbf{R}^n$); (2) $pf \to 0$ at infinity for any $p \in \mathcal{P}(\mathbf{R}^n)$. The implication (2) $\Rightarrow$ (1) is clear. To see that (1) $\Rightarrow$ (2) define q on $\mathbf{R}^n$ by $q(x) = x_1^2 + x_2^2 + \cdots + x_n^2$, $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$; assuming (1), pqf is bounded for any $p \in \mathcal{P}(\mathbf{R}^n)$, $q \to \infty$ at infinity, hence $pf \to 0$ at infinity.

Bernstein problem (first form). Let $\omega \geq 0$ be upper-semicontinuous on $\mathbf{R}^n$ and rapidly decreasing at infinity, i.e., $\mathcal{P}(\mathbf{R}^n) \subset C\omega_\infty(\mathbf{R}^n)$, or equivalently, $\mathcal{P}(\mathbf{R}^n) \subset C\omega_b(\mathbf{R}^n)$.

The weight ω is said to be *fundamental* if $\mathcal{P}(\mathbf{R}^n)$ is dense in $C\omega_\infty(\mathbf{R}^n)$; and the Bernstein problem consists in finding necessary and sufficient conditions for a given weight ω to be fundamental.

We remark here that the Weierstrass theorem means that every characteristic function of a compact subset of $\mathbf{R}^n$ is a fundamental weight; and this implies that every $\omega \geq 0$ which is upper-semicontinuous on $\mathbf{R}^n$ and has compact support, is a fundamental weight.

Bernstein problem (second form). Let $C(\mathbf{R}^n)$, and suppose w is rapidly decreasing at infinity, i.e., $\mathcal{P}(\mathbf{R}^n) \cdot w \subset C_\infty(\mathbf{R}^n)$ or equivalently $\mathcal{P}(\mathbf{R}^n) \cdot w \subset C_b(\mathbf{R}^n)$. We then say that *the load w is fundamental* if $\mathcal{P}(\mathbf{R}^n) \cdot w$ is dense in $C_\infty(\mathbf{R}^n)$. The Bernstein problem consists in finding necessary and sufficient conditions for a given load w to be fundamental.

For convenience in the sequel, we shall call these problems *Bernstein's problem I* and *Bernstein's problem II*, respectively. The next proposition follows.

Proposition 2.7. *Let $w \in C(\mathbf{R}^n)$, $w \geq 0$. Then w is a fundamental load if and only if w is a fundamental weight and $w(x) > 0$ for any $x \in \mathbf{R}^n$.*

In order to explain the work of L. Nachbin in this area we have to explain the *"weighted approximation problem"* for modules.

Let $A \subset C(E)$ be a subalgebra containing the unit, and $W \subset CV_\infty(E)$ be a vector subspace; we shall also assume W to be a module over A i.e., $AW \subset W$. The *weighted approximation problem* consists in determining $\overline{W}$ in $CV_\infty(E)$ under these circumstances.

In the special case in which A consists only of the constant functions, W is the most

general vector subspace of $CV_\infty(E)$. In this case, all we can say about $\overline{W}$ is that $\overline{W}$ consists of all $f \in CV_\infty(E) \ni$ every $\phi \in CV_\infty(E)^*$ vanishing on W must also vanish at f. The general case of the weighted approximation problem is reduced to the special case just mentioned, by considering the subsets of E on which the functions belonging to A are constant, i.e., by introducing on E the equivalence relation E/A mentioned earlier. The following definition is formulated with this view in mind.

Definition 2.8. *We say that W localisable under A in $CV_\infty(E)$ if the following holds: $\forall f \in CV_\infty(E), f \in \overline{W}$ (in $CV_\infty(E)$) $\Leftrightarrow \forall v \in V$, $\forall \varepsilon > 0$ and $\forall$ equivalence class X modulo E/A, $\exists w \in W \ni$*

$$v(x) \cdot |w(x) - f(x)| < \varepsilon \qquad \forall x \in X\ .$$

The *strict weighted approximation problem* consists in finding necessary and sufficient conditions for W to be localisable under A in $CV_\infty(E)$.

We note that if the following conditions are satisfied:

(1) A is separating on E;

(2) W is everywhere different from 0 in E i.e., $\forall x \in E \exists w \in W \ni w(x) \neq 0$;

then W is localisable under A in $CV_\infty(E) \Leftrightarrow W$ is dense in $CV_\infty(E)$. Hence if the conditions (1) and (2) are satisfied then corresponding to every sufficient condition for localisability to be established below there will be a corollary asserting density of W in $CV_\infty(E)$.

Furthermore the strict weighted approximation problem can be seen to be a generalisation of the Bernstein approximation problem, as follows. Consider the Bernstein problem I; let $E = \mathbf{R}^n$, $V = \{\omega\}$, $A = \mathcal{P}(\mathbf{R}^n)$, $W = \mathcal{P}(\mathbf{R}^n)$; or consider the Bernstein problem

II; let $E = \mathbf{R}^n$, $V = \{1\}$, $A = \mathcal{P}(\mathbf{R}^n)$, $W = \mathcal{P}(\mathbf{R}^n) \cdot w$. Then condition (1) in the preceding paragraph is satisfied; the condition (2) is always satisfied in the case of Bernstein's problem I; and as for Bernstein's problem II, the condition (2) amounts to saying that $w(x) \neq 0$ for any $x \in \mathbf{R}^n$, and in this case Proposition 2.7 justifies assuming the condition (2). Hence if these conditions (1) and (2) hold, then finding necessary and sufficient conditions for $\overline{\mathcal{P}(\mathbf{R}^n)} = C\omega_\infty(\mathbf{R}^n)$ in Bernstein's problem I, or for $\overline{\mathcal{P}(\mathbf{R}^n)w} = C_\infty(\mathbf{R}^n)$ in Bernstein's problem II, is equivalent to finding necessary and sufficient conditions for localisability.

The next step is to consider how the weighted approximation problem can be reduced to a finite-dimensional Bernstein problem. We shall denote by Ω_n the set of all upper-semicontinuous functions $\omega \geq 0$ on $\mathbf{R}^n$ which are fundamental weights in the sense of Bernstein.

Let $G(A)$ be a subset of A which topologically generates A as an algebra over $\mathbf{C}$ with unit i.e., $\ni$ the subalgebra over $\mathbf{C}$ of A generated by $G(A)$ and 1 is dense in A (in the topology of $C(E)$); also let $G(W)$ be a subset of $W \ni G(W)$ generates W as a module over A i.e., the submodule over A of W, generated by $G(W)$ is dense in W for the topology of $CV_\infty(E)$.

The following theorem now holds.

Theorem 2.9. *Suppose $C(E) = C(E;\mathbf{R})$; if we let $C(E) = C(E;\mathbf{C})$ then we shall assume that $G(A)$ consists of real-valued functions. Suppose further that $\forall v \in V, \forall a_1, \ldots, a_n \in G(A)$ and $\forall w \in G(W)$, $\exists a_{n+1}, \ldots, a_N \in G(A)$ with $N \geq n$, and $\exists \omega \in \Omega_N \ni v(x) \cdot |w(x)| \leq \omega\big(a_1(x), \ldots, a_n(x), \ldots, a_N(x)\big)$, for any $x \in E$. Then W is*

localisable under A in $CV_\infty(E)$.

For the proof we shall need the following two lemmas.

Lemma 2.10. *Let $E = \Pi_{i\in I} E_i$ be a Cartesian product of Hausdorff spaces and $\mathcal{K}$ a collection of compact subsets E with an empty intersection. Then $\exists$ finite subset $J \subset I \ni$ if Π_J denotes the natural projection $E \to \Pi_{i\in J} E_i$ then $\Pi_J(\mathcal{K})$ has an empty intersection.*

Lemma 2.11. *Let $f \in CV_\infty(E)$, $v \in V$, and $\varepsilon > 0$. Further suppose $\forall$ equivalence class $X \subset E$ modulo E/A $\exists w \in W \ni v(X)|w(x) - f(x)| < \varepsilon \forall x \in X$. Then $\exists a_1, \ldots, a_n \in G(A), \exists w_1, \ldots w_m \in G(W)$ and $\alpha_1, \ldots, \alpha_m \in C_b(\mathbf{R}^n) \ni$*

$$v(x) \cdot \left| \sum_{i=1}^{m} \alpha_i\big(a_1(x), \ldots, a_n(x)\big) w_i(x) - f(x) \right| \leq \varepsilon \qquad \forall x \in E .$$

Proof. Let F denote the space of all real-valued functions on $G(A)$, and we shall assume that F is endowed with the Cartesian product topology. Let $\pi : E \to F$ be the continuous mapping which to $x \in E$ associates $\pi(x) \in F \ni$ the value of $\pi(x)$ at $a \in G(A)$ is $a(x) \in \mathbf{R}$. Let $y \in \pi(E)$, then every $a \in G(A)$ is constant on $\pi^{-1}(y)$. Now $G(A)$ topologically generates A, hence every $a \in A$ is constant on $\pi^{-1}(y)$. Therefore $\pi^{-1}(y)$ is contained in an equivalence class modulo $E/A \cdot \pi^{-1}(y)$ is actually an equivalence class, for π is constant on every equivalence class.

Now by the assumption of the lemma, for each $y \in \pi(E)$ $\exists w_y \in W \ni v(x) \cdot |w_y(x) - f(x)| < \varepsilon$ $\forall x \in \pi^{-1}(y)$. We shall assume w_y is in the vector subspace of W generated by $G(W)$; for let

$$\delta = \sup\big\{v(x) \cdot |w_y(x) - f(x)| \;\mid\; x \in \pi^{-1}(y)\big\} ;$$

then $\delta < \varepsilon$, for $v(x) \cdot |w_y(x) - f(x)|$ attains a maximum on $\pi^{-1}(y)$. $G(W)$ topologically generates W, hence $\exists a_1, \ldots, a_r \in A$ and $\exists w_1, \ldots, w_r \in G(W) \ni$

$$v(x) \cdot |a_1(x)w_1(x) + \cdots + a_r(x)w_r(x) - w_y(x)| < \varepsilon - \delta \qquad \forall x \in E \ .$$

Let $\lambda_1, \ldots, \lambda_r$ respectively, be the constant values of $a_1, \ldots, a_r$ on the equivalence class $\pi^{-1}(y)$. Then

$$\begin{aligned} & v(x) \cdot |\lambda_1 w_1(x) + \cdots + \lambda_r w_r(x) - f(x)| \\ & \leq \ v(x) \cdot |a_1(x)w_1(x) + \cdots + a_r(x)w_r(x) - w_y(x)| \\ & + \ v(x) \cdot |w_y(x) - f(x)| \ < \ (\varepsilon - \delta) + \delta = \varepsilon \qquad \forall x \in \pi^{-1}(y) \ . \end{aligned}$$

We can replace w_y by $\sum \lambda_i w_i$, hence we may assume that w_y belongs to the vector subspace of W generated by $G(W)$. Now let

$$K_y \ = \ \{x \in E \mid v(x) \cdot |w_y(x) - f(x)| \geq \varepsilon\}$$

$\forall y \in \pi(E)$. Then $K_y \cap \pi^{-1}(y) = \emptyset$, hence $\pi(K_y)$ is compact and $y \notin \pi(K_y)$. Therefore $\bigcap_{y \in \pi(E)} \pi(K_y) = \emptyset$. Now, using Lemma 2.10, we conclude that $\exists a_1, \ldots, a_n \in G(A) \ni$ if Φ is the mapping $t \to \big(a_1(t), \ldots, a_n(t)\big)$ from $E \to \mathbf{R}^n$ then $\bigcap_{y \in \pi(E)} \Phi(K_y) = \emptyset$. By the finite intersection property and compactness, $\exists y_1, \ldots, y_p \in \pi(E) \ni \Phi(K_{y_1}) \cap \ldots \cap \Phi(K_{y_p}) = \emptyset$. Then, $\mathbf{R}^n$ being normal, $\exists$ continuous partition of unity $\beta_1, \ldots, \beta_p \in C(\mathbf{R}^n) \ni \beta_1 + \cdots + \beta_p = 1$ and $\beta_i = 0$ on $\phi(K_{y_i})$ $i = 1, \ldots, p$.

A little argument now shows that

$$v(x) \cdot \left| \sum_{i=1}^{p} \beta_i\big(a_1(x), \ldots, a_n(x)\big) w_{y_i}(x) - f(x) \right| \leq \varepsilon \ . \tag{3}$$

Finally each w_{y_i} belongs to the vector subspace of W generated by $G(W)$, hence $\exists w_1, \ldots, w_m \in G(W) \ni$ each w_{y_i} is a linear combination of $w_1, \ldots, w_m$. Then for suitable linear conbinations $\alpha_1, \ldots, \alpha_m$ of $\beta_1, \ldots, \beta_p$, the lemma follows. Note that $\alpha_i \in C_b(\mathbf{R}^n)$ for $i = 1, \ldots, m$ because $0 \leq \beta_i \leq 1$ for $i = 1, \ldots, p$.

Proof of Theorem 2.9. One preliminary remark would be in order here. If $v \in V$, $a_1, \ldots, a_n \in G(A)$, $w \in G(W)$, $\alpha \in C_b(\mathbf{R}^n)$ and $\delta > 0$, then $\exists w' \in W \ni$

$$v(x) \cdot \mid w'(x) - \alpha\big(a_1(x), \ldots, a_n(x)\big) w(x) \mid < \delta \quad \forall x \in E\ . \tag{4}$$

To justify this we note that by assumption $\exists a_{n+1}, \ldots, a_N \in G(A)$ where $N \geq n$ and $\exists \omega \in \Omega_N \ni$

$$v(x) \cdot |w(x)| \leq \omega\big(a_1(x), \ldots, a_N(x)\big) \tag{5}$$

for any $x \in E$. Now $\alpha \in C_b(\mathbf{R}^n)$ determines $\beta \in C_b(\mathbf{R}^N)$ by the formula $\beta(t_1, \ldots t_n, \ldots, t_N) = \alpha(t_1, \ldots, t_n)$ for $t_1, \ldots, t_N \in \mathbf{R}$. Also $C_b(\mathbf{R}^N) \subset C\omega_\infty(\mathbf{R}^N)$ since $\omega \to 0$ at infinity. Now because $\mathcal{P}(\mathbf{R}^N)$ is dense in $C\omega_\infty(\mathbf{R}^N)$ and $\beta \in C\omega_\infty(\mathbf{R}^N)$, therefore $\exists p \in \mathcal{P}(\mathbf{R}^N) \ni$

$$\omega(t_1, \ldots, t_N) \cdot |p(t_1, \ldots, t_n, \ldots, t_N) - \alpha(t_1, \ldots, t_n)| < \delta \tag{6}$$

for any $t_1, \ldots, t_n, \ldots, t_N \in \mathbf{R}$. Hence by (5) and (6)

$$\begin{aligned} &v(x) \cdot \mid p\big(a_1(x), \ldots, a_n(x), \ldots, a_N(x)\big) w(x) - \alpha\big(a_1(x), \ldots, a_n(x)\big) w(x) \mid \\ &\leq \omega\big(a_1(x), \ldots, a_n(x), \ldots, a_N(x)\big) \cdot \mid p\big(a_1(x), \ldots, a_n(x), \ldots, a_N(x)\big) \\ &\quad - \alpha\big(a_1(x), \ldots, a_n(x)\big) \mid < \delta \end{aligned}$$

for any $x \in E$, which proves (4).

Now to complete the proof of Theorem 2.9, let $f \in CV_\infty(E) \ni \forall v \in V$, $\forall \varepsilon > 0$ and $\forall$ equivalence class $X \subset E$ modulo E/A, $\exists w \in W \ni v(x) \cdot |w(x) - f(x)| < \varepsilon \ \forall x \in X$. Then by Lemma 2.11, $\exists a_1, \ldots, a_n \in G(A)$, $\exists w_1, \ldots, w_m \in G(W)$ and $\exists \alpha_1, \ldots, \alpha_m \in C_b(\mathbf{R}^n) \ni$

$$v(x) \cdot \left| \sum_{i=1}^{m} \alpha_i\big(a_1(x), \ldots, a_n(x)\big) w_i(x) - f(x) \right| \leq \varepsilon \qquad \forall x \in E\,. \tag{7}$$

By the preliminary remark in the last paragraph $\exists w_1', \ldots, w_m' \in W \ni$

$$v(x) \cdot \mid w_i'(x) - \alpha_i\big(a_1(x), \ldots, a_n(x)\big) w_i(x) \mid < \delta \tag{8}$$

for any $x \in E$, and $i = 1, \ldots, m$. Putting together (7) and (8), and taking $\delta = \frac{\varepsilon}{m}$, we find

$$v(x) \cdot \mid w(x) - f(x) \mid < 2\varepsilon \qquad \forall x \in E\,.$$

Hence $f \in \overline{W}$. This proves localisability of W under A in $CV_\infty(E)$.

Corollary 2.12. *Suppose $C(E) = C(E; \mathbf{R})$; or suppose $C(E) = C(E; \mathbf{C})$ and that $G(A)$ consists of real functions. Suppose further that $G(A), G(W)$ are both finite: $G(A) = \{a_1, \ldots, a_n\}$, $G(W) = \{w_1, \ldots, w_m\}$; and that $\forall v \in V, \forall i = 1, \ldots, m \ \exists \omega \in \Omega_n \ni v(x) \cdot \mid w_i(x) \mid \leq \omega\big(a_1(x), \ldots, a_n(x)\big) \forall x \in E$. Then W is localisable under A in $CV_\infty(E)$.*

Before stating the next theorem and its corollary, it is necessary to explain some notation. If $x = (x_1, \ldots x_n) \in \mathbf{R}^n$, we shall denote by $|x|$ the point $(|x_1|, \ldots, |x_n|)$. We shall also denote by Ω_n^d the set of all $\omega \in \Omega_n$ which are decreasing in the sense that if $x, y \in \mathbf{R}^n$ and $|x| \leq |y|$ then $\omega(x) \geq \omega(y)$; this implies $\omega(x) = \omega(|x|)$.

Theorem 2.13. *Suppose $C(E) = C(E; \mathbf{R})$; or suppose $C(E) = C(E; \mathbf{C})$ and that A is self-adjoint. Also suppose that $\forall v \in V$, $\forall a_1, \ldots, a_n \in G(A)$ and $\forall w \in G(W)$*

$\exists a_{n+1}, \ldots, a_N \in G(A)$ with $N \geq n$ and $\exists \omega \in \Omega_N^d$ satisfying

$$v(x) \cdot |w(x)| \leq \omega\big(|a_1(x)|, \ldots, |a_n(x)|, \ldots, |a_N(x)|\big) \qquad \forall x \in E \ .$$

Then W is localisable under A in $CV_\infty(E)$.

Remarks (1) We note that corresponding to Corollary 2.12, there is an analogous corollary of Theorem 2.13.

(3) Using the weighted Dieudonné's Theorem 2.6 on density in tensor products, the arguments of Theorems 2.9 or 2.13 are reduced to one-dimensional arguments.

We shall denote by Γ_n the set of all upper semi-continuous $\gamma \geq 0$ on $\mathbf{R}^n \ni \gamma^k$ is a fundamental weight in the sense of Bernstein for any $k > 0$. By taking $k = 1$ we see that $\Gamma_n \subset \Omega$; however there are examples showing that $\Gamma_n \neq \Omega_n$. Also if $\gamma^k \in \Omega_n$ for a certain $k > 0$ then $\gamma^\ell \in \Omega_n$ for all $\ell > k$. Furthermore we shall denote by Γ_n^d the set of all $\gamma \in \Gamma_n$ which are decreasing i.e., $x, y \in \mathbf{R}^n, |x| \leq |y| \Rightarrow \gamma(x) \geq \gamma(y)$ hence $\gamma(x) = \gamma(|x|)$.

Theorem 2.14. *Suppose A is self-adjoint and that $\forall v \in V, \forall a \in G(A)$ and $\forall w \in G(W)$ $\exists \gamma \in \Gamma_1^d \ni v(x) \cdot |w(x)| \leq \gamma(|a(x)|) \forall x \in E$. Then W is localisable under A in $CV_\infty(E)$.*

§3. Criteria of localisability

In this section several criteria of localisability due to Nachbin, will be established; we find that each of these turns out to be a special case of the one immediately following.

Theorem 3.1. *Suppose $C(E) = C(E;\mathbf{R})$; or suppose $C(E) = C(E;\mathbf{C})$ and that A is self-adjoint. In either case, suppose further that $\forall v \in V$, $\forall a \in G(A)$ and $\forall w \in G(W) \exists C > 0$ and $\exists c > 0$ satisfying*

$$v(x)|w(x)| \leq Ce^{-c|a(x)|} \qquad \forall x \in E.$$

Then W is localisable under A in $CV_\infty(E)$.

For the proof we need the following two lemmas.

Lemma 3.2. *Suppose V is a directed set of upper-semi continuous positive functions of E. Suppose $C_b(E) \subset CV_\infty(E)$ and that A is a subalgebra of $C_b(E)$ which is separating on E, contains the constant function 1 and is self-adjoint in the complex case. Then A is dense in $CV_\infty(E)$.*

Lemma 3.3. *Let $\gamma \geq 0$ upper-semicontinuous on $\mathbf{R}$ be $\ni \exists C > 0, \exists c > 0$ satisfying:*

$$\gamma(t) \leq Ce^{c|t|} \, \forall \, t \in \mathbf{R}.$$

Then $\gamma \in \Gamma_1$.

Proof. Clearly $\gamma(\cdot)$ is rapidly decreasing at infinity, since $t^m e^{-c|t|} \to 0$ as $t \to \infty$, for any $m \in \mathbf{R}$. Let $t, x, y \in \mathbf{R}$, $z = x + iy \in \mathbf{C}$, and define $e_z \in C(\mathbf{R};\mathbf{C})$ by $e_z(t) = e^{izt}$ for

$t \in \mathbf{R}$. We then note

$$e_z \in C_b(\mathbf{R};\mathbf{C}), e_0 = 1, \bar{e}_z = e_{-\bar{z}}, \quad \text{and} \quad e_{z_1+z_2} = e_{z_1} e_{z_2} \ .$$

Now

$$\gamma(t) \mid e^{izt} | \leq C e^{-(yt+c|t|)} \to 0 \quad \text{as} \quad t \to \infty$$

provided $|y| < c$. Hence, $e_z \in C_b(\mathbf{R},\mathbf{C})$ if $|y| < c$. Denote by S the open strip:

$$\left\{ z = x + iy \in \mathbf{C} \mid x, y \in \mathbf{R}, |y| < c \right\} \ .$$

Now let ϕ be a continuous linear functional on $C\gamma_\infty(\mathbf{R};\mathbf{C})$, and define $f : S \to \mathbf{C}$ by: $f(z) = \phi(e_z), z \in S$.

Our *claim* now is: $f(\cdot)$ is analytic on S. To show this we write

$$\gamma(t) e^{izt} \ = \ \sum_{m=0}^{\infty} \frac{1}{m!} \gamma(t)(izt)^m \tag{1}$$

for $z \in \mathbf{C}$, $t \in \mathbf{R}$. We shall show that if $z \in \mathbf{C}$ and $|z| < c$, then the above series (1) converges uniformly for $t \in \mathbf{R}$. We use the elementary remark that if $m > 0$, is an integer and if $c > 0$, then the function $t^m e^{-ct}$ on $t \geq 0$ attains its maximum at $t = \frac{m}{c}$; and hence

$$t^m e^{-ct} \leq \left(\frac{m}{ce}\right)^m \quad \text{for} \quad m > 0, c > 0, t \geq 0 \ .$$

Therefore for $m > 0$

$$\mid \frac{1}{m!} \gamma(t)(izt)^m \mid \leq \frac{C}{m!} \left(\frac{m|z|}{ce}\right)^m \ . \tag{2}$$

Now Stirling's formula

$$\lim_{m \to \infty} \frac{m!}{m^m e^{-m} \sqrt{2\pi m}} \ = \ 1$$

implies that

$$\lim_{m\to\infty} \frac{m}{(m!)^{1/m}} = e .$$

Hence using Cauchy's criterion for convergence of a series of positive terms, we see that

$$\sum_{m=0}^{\infty} \frac{C}{m!}\left(\frac{m|z|}{ce}\right)^m < \infty \quad \text{if} \quad |z| < c . \tag{3}$$

Thus from (2) and (3) we see that the series in (1) converges uniformly for $t \in \mathbf{R}$ and $z \in \mathbf{C}$ with $|z| < c$.

Now define $u_m \in C(\mathbf{R}; \mathbf{C})$ by $u_m(t) = (it)^m$ for $t \in \mathbf{R}$, $m = 0, 1, \ldots$; then $u_m \in C\gamma_\infty(\mathbf{R}; \mathbf{C})$ because $t^m e^{-c|t|} \to 0$ as $t \to \infty$. Because of the uniform convergence of (1) for $t \in \mathbf{R}$ we see that

$$e_z = \sum_{m=0}^{\infty} \frac{1}{m!} u_m z^m \quad \text{if} \quad |z| < c ,$$

where convergence is understood in the sense of $C\gamma_\infty(\mathbf{R}; \mathbf{C})$. Now let $a \in \mathbf{R}$: then we also have

$$e_z = e_a e_{z-a} = \sum_{m=0}^{\infty} \frac{1}{m!} e_a u_m (z-a)^m \quad \text{if} \quad |z - a| < c ,$$

where convergence is understood in the sense of $C\gamma_\infty(\mathbf{R}; \mathbf{C})$, hence $e_a u_m \in C_{\gamma_\infty}(\mathbf{R}; \mathbf{C})$. Thus we find

$$f(z) = \phi(e_z) = \sum_{m=0}^{\infty} \frac{1}{m!} \phi(e_a u_m)(z-a)^m$$

provided $|z - a| < c$. Here $a \in \mathbf{R}$ is arbitrary; thus f is seen to be analytic on S. Furthermore

$$f^{(m)}(a) = \phi(e_a u_m), \quad a \in \mathbf{R}, m = 0, 1, 2, \ldots . \tag{4}$$

Now suppose ϕ vanishes on $P(\mathbf{R}; \mathbf{C})$. Then $f^{(m)}(0) = 0$ for $m = 0, 1, 2, \ldots$ by (4) (taking $a = 0$). Since f is analytic on S, it follows that f is identically zero on S, hence

on $\mathbf{R}$, i.e., $\phi(e_x) = 0\ \forall x \in \mathbf{R}$. Denote by A the vector subspace of $C_b(\mathbf{R};\mathbf{C})$ generated by all $e_x, x \in \mathbf{R}$. Then $\phi = 0$ on A.

It is then clear that A is a self-adjoint separating subalgebra of $C_b(\mathbf{R};\mathbf{C})$ containing 1 and $C_b(\mathbf{R};\mathbf{C}) \subset C\gamma_\infty(\mathbf{R};\mathbf{C})$. By the last lemma A is dense in $C\gamma_\infty(\mathbf{R};\mathbf{C})$, hence $\phi = 0$ on $C\gamma_\infty(\mathbf{R};\mathbf{C})$. Hence, we see that

$$\phi = 0 \quad \text{on} \quad \mathcal{P}(\mathbf{R};\mathbf{C}) \Rightarrow \phi = 0 \quad \text{on} \quad C\gamma_\infty(\mathbf{R};\mathbf{C}) \; .$$

Hence by the Hahn-Banach theorem we see that $\mathcal{P}(\mathbf{R};\mathbf{C})$ is dense in $C\gamma_\infty(\mathbf{R};\mathbf{C})$. Hence $\gamma \in \Omega_1$. But now γ^k satisfies the same kind of assumption as γ for integer $k > 0$, hence $\gamma^k \in \Omega$, for $k > 0$ hence $\gamma \in \Gamma_1$. This completes the proof of the Lemma.

Proof of Theorem 3.1. We now apply Theorem 2.14, and the last Lemma, taking $\gamma(t) = Ce^{-c|t|}$ for $t \in \mathbf{R}$, where we notice that $\gamma \in \Gamma_1^d$. This proves the theorem.

We shall next turn to the quasi-analytical criterion of localisability. The following theorem will be established.

Theorem 3.4. *We shall suppose A is self-adjoint, and also that $\forall v \in V, \forall a \in G(A)$ and $\forall w \in G(W)$ we have*

$$\sum_{m=1}^{\infty} \frac{1}{M_m^{1/m}} = \infty$$

where $M_m = sup\{v(x) \cdot \mid a(x)^m w(x) \mid \,\big|\, x \in E\}$ for $m = 0, 1, 2, \ldots$. Then W is localisable under A in $CV_\infty(E)$.

Before turning to the proof of Theorem 3.4, we recall the following concepts from the area of infinitely differentiable functions. Suppose $\mathbf{M} = \{M_m \mid m = 0, 1, 2, \ldots\}$ is a sequence of strictly positive numbers. We shall denote by $C(\mathbf{M})$ the set of all indefinitely

differentiable complex-valued functions f, each defined on some open interval $I \subset \mathbf{R}$ (I depending on f), and satisfying the following estimates for its successive derivatives: for every compact subset $K \subset I$, $\exists C > 0$, and $\exists c > 0 \ni$

$$| f^{(m)}(x) | \leq C \cdot c^m \cdot M_m$$

$\forall x \in K$ and $m = 0, 1, 2, \ldots$. We say that $C(\mathsf{M})$ is a *quasi-analytic* class if the following is true: if $f \in C(\mathsf{M})$ and $\exists a \in I$ such that $f^{(m)}(a) = 0$

$$\forall m = 0, 1, 2, \ldots, \quad \text{then} \quad f \equiv 0 \quad \text{on} \quad I.$$

Clearly this amounts to the requirement that every $f \in C(\mathsf{M})$ is determined within $C(\mathsf{M})$ by the knowledge of its Taylor series at a single point $a \in I$ (though the Taylor series of f at a or at any other point in I are not assumed to be convergent – in fact may fail to be convergent).

If $M_m = m!(m = 0, 1, 2, \ldots)$ then by a theorem of Pringsheim, $C(\mathsf{M})$ consists of complex valued functions which are analytic on open intervals of $\mathbf{R}$, and hence can be called the *analytic* class. In this case $C(\mathsf{M})$ is clearly quasi-analytic. On the other hand it is known that not every class is quasi-analytic. For instance, if

$$M_m = \sup\{|\frac{d^m}{dx^m} e^{-\frac{1}{x}}| \; \Big| \; 0 < x < 1\}$$

for $m = 0, 1, 2, \ldots$, then $C(\mathsf{M})$ is not quasi-analytic, for the function f defined by: $f(x) = e^{-\frac{1}{x}}, 0 < x < 1$, and $f(x) = 0$ for $-1 < x \leq 0$, is indefinitely differentiable on the interval $I = (-1, 1)$; furthermore, $f \in C(\mathsf{M})$, and $f^{(m)}(0) = 0$ for $m = 0, 1, 2, \ldots$, yet f is not the zero function on I.

We shall assume Denjoy's criterion of quasi-analyticity. This is contained in Theorem 3.5 below, by which Denjoy solved the problem of Hadamard, to find necessary and sufficient conditions on a given sequence M in order that $C(\mathsf{M})$ should be quasi-analytic.

Theorem 3.5. *(Denjoy-Carleman) Let M be given and set*

$$\mu_m = \inf\{M_k^{\frac{1}{k}}, k = m, m+1, \ldots\}$$

for $m = 1, 2, \ldots$. Then $C(\mathsf{M})$ is a quasi-analytic class $\Leftrightarrow \sum_{m=1}^{\infty} \frac{1}{\mu_m} = \infty$.

Corollary 3.6. *If $\sum_{m=1}^{\infty} \frac{1}{M_m^{\frac{1}{m}}} = \infty$ then $C(\mathsf{M})$ is quasi-analytic.*

We note that the Corollary follows from Theorem 3.5.

Remark. If $M_m = m!, m = 0, 1, 2, \ldots$, then the class $C(\mathsf{M})$ is the analytic class, as noted above. Since (by Stirling's formula) $\exists \lambda > 0 \ni m! \leq \lambda^m m^m (m = 1, 2, \ldots)$, it follows that $C(\mathsf{M})$ is also quasi-analytic. On the other hand, suppose $C(\mathsf{M})$ is quasi-analytic by virtue of the existence of some $\alpha > 0 \ni \frac{1}{M_m^{\frac{1}{m}}} \geq \frac{\alpha}{m} (m = 1, 2, \ldots)$ and by application of the last Corollary. Since $\exists \lambda > 0 \ni m^m \leq \lambda^m \cdot m! (m = 1, 2, \ldots)$, it follows that $C(\mathsf{M})$ is then contained in the analytic class. We conclude that the analytic class is the largest quasi-analytic class which is tied up with the divergence of the harmonic series $\sum_{m=1}^{\infty} \frac{1}{m}$.

We shall need two lemmas before we turn to the proof of Theorem 3.4.

Lemma 3.7. *Let $\alpha_m \geq 0 (m = 1, 2, \ldots)$ and suppose $\exists \sigma > 0 \ni \alpha_{m+1} \leq \sigma \alpha_m$ for $m = 1, 2, \ldots$ If $\sum_{m=1}^{\infty} \alpha_m = \infty$ then $\sum_{m=1}^{\infty} \alpha_{mp} = \infty$ for $p = 1, 2, \ldots$ The formal proof will be omitted.*

Lemma 3.8. *Let $\gamma \geq 0$ upper-semicontinuous $\ni$*

$$\sum \frac{1}{M_m^{\frac{1}{m}}} = \infty$$

where $M_m = \sup\{\gamma(t).\quad |t^m| | t \in \mathbf{R}\}$ *for* $m = 0, 1, 2, \ldots$. *Then* γ *is fundamental weight on* $\mathbf{R}$. *To be precise,* $\gamma \in \Gamma_1$.

Proof of Lemma 3.8. If $M_m = 0$, for some m, then $\gamma(t) = 0$ for $t \neq 0$. The support of $\gamma(\cdot)$ is reduced to 0, or empty, hence by an earlier remark (re: the Bernstein Approximation problem) we see that $\gamma \in \Gamma_1$.

Next suppose $M_m > 0$ for $m = 0, 1, 2, \ldots$. Since the series $\sum_{m=1}^{\infty} \frac{1}{M_m^{\frac{1}{m}}}$ is assumed to diverge, it follows that $M_m < \infty$ for infinitely many values of m. On the other hand, if $M_m < \infty$ for some m, then $M_p < \infty$ for all $p = 0, \ldots, m$, as we find from the relation:

$$\gamma(t)|t|^p = \frac{\gamma(t)|t^m|}{|t^{m-p}|}$$

outside a compact neighbourhood of 0. Thus γ is rapidly decreasing at infinity.

Using the earlier notation for e_x and u_m, if $x \in \mathbf{R}$, then $e_x \in C_b(\mathbf{R}; \mathbf{C}) \subset C\gamma_\infty(\mathbf{R}; \mathbf{C})$. Let ϕ be a continuous linear functional on $C\gamma_\infty(\mathbf{R}; \mathbf{C})$. Define $f : \mathbf{R} \to \mathbf{C}$ as before by $f(x) = \phi(e_x)$, $x \in \mathbf{R}$. We then *claim* that f is C^∞ on $\mathbf{R}$, and

$$f^{(m)}(a) = \phi(e_a u_m) \tag{5}$$

$a \in \mathbf{R}, m = 0, 1, 2, \ldots$ (here we note that $e_a \in C_b(\mathbf{R}; \mathbf{C})$ and $u_m \in C\gamma_\infty(\mathbf{R}; \mathbf{C})$, hence $e_a u_m \in C\gamma_\infty(\mathbf{R}; \mathbf{C})$). For suppose f is m-differentiable and (5) is true for some $m \geq 0$. This is true for $m = 0$. If $h \in \mathbf{R}$, $h \neq 0$, then

$$\frac{f^{(m)}(a+h) - f^{(m)}(a)}{h} = \phi\big(e_a \cdot \frac{e_h - 1}{h} \cdot u_m\big). \tag{6}$$

By Taylor's formula, if g is twice differentiable on the interval $I(h) = [0, h]$ (or $[h, 0]$), then

$$|g(h) - g(0) - hg'(0)| \leq \frac{h^2}{2} \sup\{|g''(x)| | x \in I(h)\}.$$

Apply this to $g(x) = e^{ixt}, t \in \mathbf{R}$ being fixed; then we obtain

$$|e^{iht} - 1 - iht| \leq \frac{(ht)^2}{2}.$$

Hence,

$$\gamma(t) \left| e^{iat}\Big(\frac{e^{iht} - 1}{h} - it\Big)(it)^m \right| \leq \frac{hM_{m+2}}{2}.$$

Therefore,

$$e_a \cdot \frac{e_h - 1}{h} \cdot u_m \rightarrow e_a \cdot u_{m+1} \tag{7}$$

as $h \rightarrow 0$, the convergence being in $C\gamma_\infty(\mathbf{R}; \mathbf{C})$. Then from (6) and (7) it follows that $f^{(m)}$ is differentiable and (5) holds with m replaced by $m+1$. This proves our claim (5).

From (1) it follows that

$$|f^{(m)}(a)| \leq ||\phi|| M_m \qquad (a \in \mathbf{R}, \qquad m = 0, 1, 2, \dots)$$

where $||\phi||$ is the seminorm of ϕ on the semi-normed space $C\gamma_\infty(\mathbf{R}; \mathbf{C})$. Hence $f \in C(\mathbf{M})$ with $\mathbf{M} = \{M_m, m = 0, 1, 2, \dots\}$. By Denjoy's Corollary above, $C(\mathbf{M})$ is quasi-analytic.

Now suppose $\phi = 0$ on $\mathcal{P}(\mathbf{R}; \mathbf{C})$. Then $f^{(m)}(0) = 0$, for $m = 0, 1, 2, \dots$ by (5) for $a = 0$. Hence by quasi-analyticity, $f \equiv 0$ on $\mathbf{R}$. The rest of the proof that $\gamma \in \Gamma_1$ proceeds along the same lines as in Lemma 3.3 above.

Next, γ^k satisfies the same assumption as γ for any $k > 0$, i.e. if we define

$$N_m(k) = sup\{\gamma(t)^k |t|^m | t \in \mathbf{R}\}$$

for $m = 0, 1, 2, \dots$, then an inductive argument combined with the hypothesis of the Lemma and an application of the last Lemma shows that

$$\sum_{m=1}^{\infty} \frac{1}{N_m(k)^{\frac{1}{m}}} = \infty. \tag{8}$$

From the earlier part of the proof we then conclude that $\gamma^k \in \Omega_1$ for any $k > 0$, i.e. $\gamma \in \Gamma_1$. This proves the Lemma.

We now turn to the proof of Theorem 3.4.

Proof of Theorem 3.4. Define γ on $\mathbf{R}$ by

$$\gamma(t) = \inf\left\{ \frac{M_m}{|t^m|} \middle| m = 0, 1, 2, \dots \right\}, \quad t \in \mathbf{R},$$

where we understand that $\gamma(0) = 0$ if some $M_m = 0$, and $\gamma(0) = M_0$ otherwise. Then $\gamma \geq 0$ and γ is upper-semi-continuous, being an infimum of a family of continuous functions. By the definition of γ,

$$sup\left\{ \gamma(t)|t|^m \middle| t \in \mathbf{R} \right\} \leq M_m$$

for $m = 0, 1, 2, \dots$, hence $\gamma \in \Gamma_1$, by the last Lemma. Then clearly $\gamma \in \Gamma_1^d$. By the definition of M_m

$$v(x) \cdot |a(x)^m w(x)| \leq M_m, \; m = 0, 1, 2, \dots ,$$

hence $v(x) \cdot |w(x)| \leq \gamma(|a(x)|) \; \forall x \in E$. Now we apply Theorem 2.14. This proves Theorem 3.4.

Distinguishing between the bounded analytic and quasi-analytic cases of the weighted approximation problem

By "the *bounded case* of the weighted approximation problem" is meant the one in which every $a \in G(A)$ is bounded on the support of every vw, for $v \in V, w \in G(W)$. We note that each of the following assumptions leads to an instance of the bounded case:

1. $A \subset C_b(E)$;
2. every $a \in G(A)$ is bounded on the support of every $v \in V$;
3. every $a \in G(A)$ is bounded on the support of every $w \in G(W)$;
4. each vw, for $v \in V$, and $w \in G(W)$, has compact support;
5. each $v \in V$ has compact support;
6. each $w \in G(W)$ has compact support.

The next few propositions in this section are meant to distinguish between the different cases of localisability.

Proposition 3.9. *The bounded case arises $\Leftrightarrow$ in Theorem 3.4, γ can be taken to have compact support; or equivalently, if γ can be taken to be a constant times the characteristic function of a compact subset of* $\mathbf{R}$.

Proposition 3.10. *The bounded case arises $\Leftrightarrow$ in Theorem 3.4 the series is divergent because $\exists c > 0 \ni \frac{1}{M_m^{\frac{1}{m}}} \geq c$ for $m = 1, 2, \ldots$*

Proof. Suppose $\exists c > 0 \ni \frac{1}{M_m^{\frac{1}{m}}} \geq c$. Then $|v(x)w(x)|^{\frac{1}{m}} \cdot |a(x)| \leq \frac{1}{c}$ for $m = 1, 2, \ldots,$ and $x \in E$. If $v(x)w(x) \neq 0$, let $m \to \infty$, and we obtain $|a(x)| \leq \frac{1}{c}$. Hence $a(x)$ is bounded on the set $\{x \in E \mid v(x)w(x) \neq 0\}$, hence on its closure, i.e. the support of vW. Conversely suppose $|a| \leq k$, $k > 0$, on the support of vw. Suppose vw is bounded

by $C > 0$ on E. We note that $w \in CV_\infty(E) \subset CV_b(E)$ and $v \in V$. Then $\mid v(x)w(x) \mid^{1/m}$ $\mid a(x) \mid \leq C^{1/m}k$ for $m = 1, 2, \ldots$ and $x \in E$. If we choose $c \ni 0 < c < \frac{C^{-1/m}}{k}$ for $m = 1, 2, \ldots$, then we shall have $\frac{1}{M_m^{1/m}} \geq c$ for $m = 1, 2, \ldots,$. This proves this proposition.

Theorem 3.11. *With the preceding definitions and terminology, localisability always holds in the bounded case of the weighted approximation problem, provided we deal with real valued functions, or complex-valued functions and A is assumed to be self-adjoint.*

Proof of Theorem 3.11. Use the first of last two propositions and Theorem 2.14.

By the *analytic case* or the *quasi-analytic case* of the weighted approximation problem is meant the one in which the sufficient condition of Theorem 3.1, or of Theorem 3.4, holds. This terminology is justified by the fact that in the proof of Theorem 3.1 or of Theorem 3.4 (respectively) use was made of analyticity (or of quasi-analyticity).

The next proposition distinguishes the occurrence of the analytic case.

Proposition 3.12. *The analytic case arises if and only if in Theorem 3.4 the series is divergent because $\exists c > 0 \ni \frac{1}{M_m^{1/m}} \geq \frac{c}{m}$ for $m = 1, 2, \ldots$.*

Proof. Suppose $\exists c > 0 \ni \frac{1}{M_m^{1/m}} \geq c$ for $m = 1, 2, \ldots$. Let $C = \sup\{1, M_0\}$. Then $c^m M_m \leq C$, i.e. $c^m v(x) \cdot |a(x)^m w(x)| \leq C$ for $m = 0, 1, 2, \ldots$, and $x \in E$. Dividing by $m!$, and summing up, we obtain

$$v(x) \cdot |w(x)| e^{c|a(x)|} \leq Ce ,$$

hence

$$v(x) \cdot |w(x)| \leq Ce \cdot e^{-c|a(x)|} \quad \text{for any} \quad x \in E .$$

Hence the sufficient condition of Theorem 10 holds.

Conversely suppose $\exists C > 0$, and $\exists c > 0 \ni$

$$v(x) \cdot |w(x)| \leq Ce \cdot e^{-c|a(x)|} \quad \forall x \in E .$$

We now note the elementary inequality

$$t^m e^{-ct} \leq \left(\frac{m}{ce}\right)^m \quad \text{for} \quad m > 0, t \geq 0$$

which has already been noted above in the course of the proof of Theorem 3.1. Then we find:

$$v(x) \cdot |a(x)^m w(x)| \leq C\left(\frac{m}{ce}\right)^m \quad \text{for any} \quad x \in E .$$

Choose $c' \ni 0 < c' \leq ceC^{-1/m}$ for $m = 1, 2, \ldots$ and then $\frac{1}{M_m^{1/m}} \geq \frac{c'}{m}$ for $m = 1, 2, \ldots$. This proves the proposition.

§4. A differentiable variant of the Stone-Weierstrass theorem

In this section and the next we shall give an account of differentiable analogues of the Stone-Weierstrass theorem for certain algebras of r-times continuously differentiable functions. We shall explain a theorem due to L. Nachbin (cf. [44]), and in the next section mention some generalisations by Aaron and Prolla.

Suppose M is a differentiable manifold of order $r \geq 1$, and dimension $n \geq 1$. Let $\mathcal{A}$ be the algebra of C^r-functions i.e., r-times continuously differentiable functions on M endowed with the topology of uniform convergence of C^r functions up to order m on compact subsets of M. Nachbin established the following theorem ([44]) p. 1550).

Theorem 4.1. *A necessary and sufficient condition for the algebra $\mathcal{A}(B)$ generated by a subset $B \subset \mathcal{A}$ to be dense in $\mathcal{A}$ is that the following conditions are satisfied:*

(1) for each $\xi \in M \exists f \in B \ni f(\xi) \neq 0$;

(2) for each pair of points $\xi, \eta \in M$, $\xi \neq \eta$, $\exists f \in B \ni f(\xi) \neq f(\eta)$;

(3) for each $\xi \in M$ and for each tangent vector $\theta \neq 0$ at ξ, $\exists f \in B \ni \frac{\partial f}{\partial \theta} \neq 0$.

Proof. Only the sufficiency needs justification. Let $K \subset M$ be compact, and W an open connected subset of M containing K and such that $\overline{W}$ is compact. For each point in $M \exists$ a function which is not identically 0 in a neighbourhood of this point; hence $\exists$ finite number of functions $f_1, \ldots, f_n \in B \ni \{f_1(x), \ldots, f_n\} \neq \{0, \ldots, 0\}$ for $x \in \overline{W}$.

Now let $\xi \in M$, and let $\theta_1 \neq 0$ a tangent vector to M at ξ. Then by hypothesis $\exists f_1 \in B \ni \frac{\partial f_1}{\partial \theta_1} \neq 0$. If $n \geq 2$, then $\exists$ tangent vector $\theta_2 \neq 0$ at $\xi \ni \frac{\partial f_1}{\partial \theta_2} = 0$. Then let $f_2 \in B \ni \frac{\partial f_2}{\partial \theta_2} \neq 0$. Then if $n \geq 3$, $\exists$ tangent vector $\theta_3 \neq 0$ at $\xi \ni \frac{\partial f_1}{\partial \theta_3} = \frac{\partial f_2}{\partial \theta_3} = 0$, etc. Thus we obtain tangent vectors $\theta_1, \ldots, \theta_n$ at ξ and functions $f_1, \ldots, f_n \in B \ni \frac{\partial f_i}{\partial \theta_i} \neq 0$

$(i \leq i \leq n)$ and $\frac{\partial f_i}{\partial \theta_j} = 0$ $(1 \leq i < j \leq n)$.

Consider the linear mapping defined on the tangent space at ξ with values in $\mathbf{R}^n$, which maps $\theta \to \left\{\frac{\partial f_1}{\partial \theta}, \dots \frac{\partial f_n}{\partial \theta}\right\}$. Each vector in $\mathbf{R}^n$ is the image of a vector $\theta = c_1\theta_1 + \cdots + c_n\theta_n$, i.e., this mapping is an isomorphism on $\mathbf{R}^n$.

The implicit function theorem shows that the mapping $x \to \{f_1(x), \dots, f_n(x)\}$ is a homeomorphism of order r (cf. Dieudonné: Foundations p. 272) of a neighbourhood of ξ in M onto an open subset in $\mathbf{R}^n$. The set $\overline{W}$ is compact, hence $\exists$ functions $g_1^i, \dots, g_n^i \in B$ and $\exists$ open subsets $V_i \subset M$ $(1 \leq i \leq b)$ covering $\overline{W} \ni$ each mapping $x \to \{g_1^i(x), \dots, g_n^i(x)\}$ is a homeomorphism of order r of V_i onto an open subset of $\mathbf{R}^n$.

Now set $f_{a+(i-1)n+j} = g_j^i$. If $\xi, \eta \in M$, $\xi \neq \eta$, then $\exists f \in B \ni f(x) \neq f(y)$ for all (x, y) in a neighbourhood of (ξ, η). The space

$$\Omega = \overline{W} \times \overline{W} - (V_1 \times V_1) \cup \ldots \cup (V_b \times V_b)$$

is compact and disjoint from the diagonal of $\overline{W} \times \overline{W}$ and thus $\exists$ functions $h_1, \dots, h_c \in B \ni$ for $(x, y) \in \Omega$ we have

$$\{h_1(x), \dots, h_c(x)\} \neq \{h_1(y), \dots, h_c(y)\} .$$

Write $f_{a+bn+i} = h_i$. Then we consider the mapping $\Phi : M \to \mathbf{R}^N$ with

$$N = a + bn + c$$

defined by the mapping $x \to \{f_1(x), \dots, f_N(x)\}$. This mapping Φ is a homeomorphism of order r of W on the submanifold $\Phi(W)$ of order $r^{\#}$ in $\mathbf{R}^N$.

Consider the inverse $\Phi^{-1} : \Phi(W) \to (W)$. Let $f \in A$. Then $f\Phi^{-1}$ is r-times continuously differentiable of order r on $\Phi(w)$. By a special case of a theorem of Whitney

(cf. [63]), $\exists$ r-times continuously differentiable function ϕ on $\mathbf{R}^N \ni \phi(z) = f(\Phi^{-1}(z))$ on $\Phi(K)$, hence $f(x) = \Phi(f_1(x), \ldots, f_N(x))$ on K. We note that $\Phi(\overline{W})$ does not contain the origin $\mathbf{R}^N$. Hence we can suppose ϕ is 0 at the origin in $\mathbf{R}^N$. Now we apply the classical Weierstrass theorem; the proof is thus completed.

Remark. The above theorem can be alternatively formulated as follows: *any proper closed subalgebra B of the algebra A is contained in a maximal closed subalgebra.*

For maximal closed subalgebra is precisely one of the following types of sets: either the set of functions vanishing at a point, or the set of functions taking the same value at 2 distinct points, or the set of functions with a zero derivative along one tangent.

§5. Further differentiable variants of the Stone-Weierstrass theorem

The "differentiable" result of Nachbin explained in the last section was generalised, first by Lesmes and Restrepo who dealt, respectively, with C^1-functions defined on a Hilbert space (Lesmes [34]) or on certain reflexive Banach spaces (Restrepo [49]); and later by Aron and Prolla (cf. [2]) who dealt with C^k mappings between Banach spaces E and F with E^* satisfying the bounded approximation property.

Here we shall briefly summarise the results of Aron and Prolla [2]. To explain these, it is first necessary to explain some notation and definitions.

Notation and definitions.

$\mathbb{N}$ denotes the set of natural numbers, $\mathbb{N}^* = \mathbb{N} \cup \{\infty\}$, and $0 \in \mathbb{N}$. E and F are *real* Banach spaces, their normed duals being denoted by E^* and F^*, respectively. For each $n \in \mathbb{N}$, $\mathcal{P}(^nE;F)$ is the space of continuous n-homogeneous polynomials from E to F, each polynomial being a composition of the form $A \circ \Delta_n$, where A is an element of the space $\mathcal{L}(^nE;F)$ of continuous n-linear mappings of $\underbrace{E \times \cdots \times E}_{n} \to F$ and Δ_n is the diagonal operator $\Delta_n : E \to \underbrace{E \times \cdots \times E}_{n}$. For $n = 0, \mathcal{L}(^\circ E;F)$ and $\mathcal{P}(^\circ E;F)$ are both identified with F. $\mathcal{P}(^nE;F)$ is a Banach space with the norm

$$\|P\| = \sup\{\|Px\| \mid \|x\| \leq 1, x \in E\}, P \in \mathcal{P}(^nE;F) .$$

The subspace $\mathcal{P}_f(^nE;F)$ of $\mathcal{P}(^nE;F)$ is generated by the collection of functions of the form $\phi^n \otimes y$ $(n \in \mathbb{N}, \phi \in E^*, y \in F)$ where $\phi^n \otimes y\ (x) = \phi^n(x)y$ for each $x \in E$. The completion of $\mathcal{P}_f(^nE;F)$ with respect to the norm of $\mathcal{P}(^nE;F)$ is denoted by $\mathcal{P}_c(^nE;F) \subset \mathcal{P}(^nE;F)$. $\mathcal{P}(E;F)$ is defined to be $= \sum_{n=0}^{\infty} \mathcal{P}(^nE;F)$.

Let U be a nonempty open subset of E. The space $C^m(U;F), m \in \mathbb{N}^*$, is the space

of all mappings $f : u \to F \ni \forall j \in \mathbf{N}^*, j \leq m$ and $x \in U$, the jth successive Fréchet derivative $D^j f(x) \in \mathcal{P}(^j E; F)$ exists and is a continuous function of $x \in U$. If the range space F is not specifically identified, it is understood that $F = \mathbf{R}^1$; so $C^m(U)$ denotes $C^m(U, \mathbf{R}^1)$, etc.

A subspace $A \subset C^m(U; F)$ is a *polynomial algebra* if $\forall g \in A$, and $\forall P \in \mathcal{P}_f(E; F)$ the composition $P \circ g \in A$. The work of Aron-Prolla shows that the Stone-Weierstrass theorem holds for polynomial algebras in $C^m(U; F)$.

A *Nachbin polynomial algebra* is a polynomial algebra $A \subset C^m(U; F)$ $(m \geq 1)$ satisfying:

(a) $\forall x \in U, \exists g \in A \ni g(x) \neq 0$;

(b) $\forall x, y \in U, x \neq y, \exists g \in A \ni g(x) \neq g(y)$;

(c) $\forall x \in U$ and $\forall v \in E, v \neq 0$, $\exists g \in \forall \ni Dg(x) \neq 0$.

When $F = \mathbf{R}^1$, a polynomial algebra is an algebra; in this case (a) and (b) are the usual conditions in the real version of the Stone-Weierstrass theorem. Nachbin (cf. §4) showed that in a finite-dimensional E, an algebra $A \subset C^m(U)$ is dense in $C^m(U)$ w.r.t. the compact open topology $\Leftrightarrow$ all three conditions (a)-(c) hold.

We note that $\mathcal{P}_f(E; F), \mathcal{P}(E; F), C^\infty(U; F)$, $C^m(U; F)$ are examples of Nachbin algebras. Suppose E has an m times continuously differentiable norm; in this case $[f \in C^m(E; F) \mid f$ has bounded support$]$ is also a Nachbin algebra.

A certain property called the Bounded approximation property is related to one condition which will be often assumed in this section. A Banach space E is said to have the *approximation property* if $\forall$ compact $K \subset E$, $\forall \varepsilon > 0$, $\exists$ operator T (depending on ε

and K) of finite rank $\ni \|Tx - x\| < \varepsilon \ \forall x \in K$; and E is said to have the *bounded approximation property* if it is further possible to choose the approximating T of finite rank $\ni \forall \varepsilon > 0$ and $\forall$ compact $K \subset E, \|T\| \leq \lambda$ where λ is independent of ε and K.

The following condition will be often assumed in this section: $\exists C \geq 1 \ni \forall$ compact $K \subset E$, $\forall$ compact $L \subset E^*$ and $\forall \varepsilon > 0$, $\exists \pi \in L(E; F)$ of finite rank satisfying:

$$\|\pi\| \leq C, \quad \|\pi(x) - x\| < \varepsilon \quad \forall x \in K \ , \tag{$*$}$$

and

$$\|\phi \circ \pi - \phi\| < \varepsilon \qquad \forall \phi \in L \ .$$

It is shown in [25] that "E^* has the bounded approximation property" is equivalent to the property (*).

Weakly uniformly continuous mappings.

We find that it is necessary to consider polynomials and functions which are weakly uniformly continuous when restricted to any bounded set.

Definition. Let E, F be real Banach spaces. A mapping $f : E \to F$ is *weakly uniformly continuously on bounded subsets of E* if $\forall$ bounded $B \subset E$, and $\forall \varepsilon > 0$ $\exists \phi_1, \ldots, \phi_k \in E'$ and $\exists \delta > 0 \ni$

$$x, y \in B \quad \text{and} \quad |\phi_i(x) - \phi_i(y)| < \delta \quad \text{for} \quad i = 1, \ldots, k \Rightarrow \|f(x) - f(y)\| < \varepsilon \ .$$

Definition. $\mathcal{P}_w({}^nE; F)$ is the subspace of $\mathcal{P}({}^mE; F)$ consisting of those m homogeneous continuous polynomials which are weakly uniformly continuous on bounded subsets of E (equivalently, on the unit ball in E).

Definition. $C_w^m(E;F)$ is the space of m-times continuously differentiable mappings $f : E \to F$ satisfying:

(a) $D^j f(x) \in \mathcal{P}_w({}^m E; F)$, for $x \in E, j \leq m$.

(b) $D^j f : E \to \mathcal{P}_w({}^m E; F)$ is weakly uniformly continuous on bounded subsets of $E (j \leq m)$.

(c) $C_w^\infty(E;F) = \bigcap_{m=0}^{\infty} C_w^m(E;F)$.

Remark:

(i) If E is reflexive then $f \in C_w^m(E;F)$ *iff* for each $j \leq m$, $D_f^j : E \to \mathcal{P}_w({}^m E; F)$ is weakly continuous on bounded subsets of E (cf. Restrepo [49] for the case $m = 1$, E reflexive).

(ii) $C_w^m(E;F)$ contains all functions of the form $g \circ T$ where T is a continuous linear operator of finite rank and $g \in C^m\big(T(E);F\big)$

(iii) $C_w^m(E;F)$ contains no non-zero function with bounded support except when $F = 0$ or dim $E < \infty$.

Definition. τ_b^m is the locally convex topology of uniform convergence of order m on bounded subsets of E, endowed upon $C_w^m(E;F)$, and is defined by all semi-norms of the form

$$\sup\Big\{\|D^j f(x)\| \Big| x \in B, j \leq m\Big\} \quad \text{for} \quad f \in C_w^m(E;F) ,$$

where B is an arbitrary bounded subset of E. Each such semi-norm is well-defined. The topology τ_b^∞ on $C_w^\infty(E;F)$ is defined in an obvious manner.

Definition. A function $f \in C^m(E;F)$ is said to to be *uniformly differentiable of order*

m if $\forall$ bounded $B \subset E$ and $\forall \varepsilon > 0\ \exists \delta > 0 \ni$ if $x \in B$, and $y \in E$ with $\|y\| < \delta$ then

$$\|f(x+y) - f(x) - Df(x)(y) - \ldots - \frac{D^m f(x)}{m!}(y)\| \leq \varepsilon \|y\|^m .$$

(*Note*: Restrepo [49] investigated uniform differentiability of order 1.)

The first main result proved in [2] is the following theorem on uniform approximation of C^m mappings up to order m on bounded subsets of E.

Theorem 5.1. *(cf. Aron and Prolla [2] p. 207) Suppose E, F are real Banach spaces with E^* having the bounded approximation property with constant C; let $m > 0$. Then a polynomial algebra $A \subset C_w^m(E; F)$ is τ_b^m-dense $\Leftrightarrow$ the following conditions hold:*

(a) A is a Nachbin polynomial algebra;

(b) $\forall$ continuous linear map $\pi : E \to E$ of finite rank, with $\|\pi\| \leq C$, and $\forall g \in A$, the composite $g \circ \pi$ belongs to the τ_b^m-closure of A.

Approximation up to order m in the compact-open topology.

Definition. Let $U \subset E$ be an open set, where E, F are real Banach spaces. Then $C_c^m(E; F)$ (for $m \in \mathbb{N}^*$) is the space of functions $f \in C^m(U; F) \ni$ for each $x \in U$ and $\forall j \leq m$, $D^j f(x) \in \mathcal{P}_w(^m E; F)$.

$C_c^m(E; F)$ is endowed with the locally convex topology of uniform convergence on compact sets of order m, defined by the family of semi-norms of the form

$$\sup\{\|D^j f(x)\| \Big| x \in K\}, f \in C_c^m(E; F)$$

where $j \leq m$, and K is an arbitrary compact subset of U.

Remark.

(i) When $m = 0$ or 1, $C_c^m(U;F) = C^m(U;F)$, and for $m > 1$, the two spaces are generally different, though in certain cases, e.g. when $E = c_0, F = \mathbb{R}^1, C_c^m(U;F) = C^m(U;F)\ \forall m \in \mathbb{N}^*$.

(ii) $C_w^m(E;F)$ is always a proper subset of $C_c^m(E;F)$ if E is infinite dimensional.

The second main result of [2] is the following theorem.

Theorem 5.2. *(cf. [2] p. 210) Suppose E, F are real Banach spaces, E^* having the bounded approximation property with constant C. Let $m \in \mathbb{N}^*$, and $U \subset E$ be a nonempty open set. A polynomial algebra $A \subset C_c^m(U;F)$ is τ_c^m-dense in $C_c^m(U;F) \Leftrightarrow$ the following conditions are satisfied:*

(a) A is a Nachbin polynomial algebra;

(b) $\forall$ continuous linear operator $\pi : E \to E$ of finite rank and $\|\pi\| \le C$, $\forall g \in A$ and $\forall$ open $V \subset U \ni \pi(V) \subset U$, the composite $g \circ (\pi\,|_V)$ belongs to the closure of $A\,|_V$ in $C_c^m(U;F)$.

This Page Intentionally Left Blank

CHAPTER II

Strong approximation in finite dimensional spaces

The concept of strong approximation appears to have originated with H. Whitney (cf. [66]), though he does not use the words "strong approximation" in his paper. In this chapter we shall present some results on strong approximation in a finite dimensional space $\mathbf{R}^n$. The first of these (Theorem 1.8 below) is Whitney's theorem on strong approximation by real analytic functions. We have presented the original proof (Lemma 6 in [66]), for we feel that this proof might suggest further possibilities (see also the Appendix by Stein in [1]). The second result in §2 of this Chapter, is a weaker result than Whitney's but is still interesting because it uses different techniques. This result on strong approximation by C^∞ functions appears to be rather commonly known (cf. Munkres [42], Hirsch [21]); however, we have attempted to be guided by the presentation in [21].

§1. Whitney's theorem on analytic approximation

We shall first explain some notation. A point in $\mathbf{R}^n$ shall be denoted either by a single variable, e.g., x, or by an ordered n-tuple of real numbers e.g., $(x_1, \ldots, x_n)$, and we shall write $x = (x_1, \ldots, x_n)$ for a point in $\mathbf{R}^n$. A *multi-index* is an ordered n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of non negative integers α_i. With each multi-index α is associated the differential operator

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} = D_1^{\alpha_1} \cdots D_n^{\alpha_n}$$

where $D_i = \frac{\partial}{\partial x_i}$; so $D^\alpha f(x)$ means $\frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}} f(x_1, \ldots, x_n)$. The order $|\alpha|$ of D^α is

defined by: $|\alpha| = \alpha_1 + \cdots + \alpha_n$; if $|\alpha| = 0$, then $D^\alpha f$ means f. Clearly $|\alpha+\beta| = |\alpha|+|\beta|$. $\alpha \pm \beta$ means $(\alpha_1 \pm \beta_1, \ldots, \alpha_n \pm \beta_n)$, and $\alpha \leq \beta$ means $\alpha_i \leq \beta_i$, $i = 1, \ldots, n$. We shall write

$$\binom{k}{l} \quad \text{for} \quad \binom{k_1}{l_1} \cdots \binom{k_n}{l_n}$$

where $k = (k_1, \ldots, k_n)$, $l = (l_1, \ldots, l_n)$ are multi-indices with $l \leq k$. The distance between two points $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbf{R}^n$ will be denoted by $d(x,y) = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{1/2} = \|x - y\|$. However, a little further on we shall allow x and y to be complex:

$$x = (x_1' + ix_1'', \ldots, x_n' \; ix_n''),$$

$$y = (y_1' + iy_1'', \ldots, y_n' + iy_n''),$$

in which case $d(x,y)^2$ shall mean $\sum_{j=1}^n \{(x_j' - y_j') + i(y_j'' - y_j'')\}^2$, where $i = \sqrt{-1}$ in $\mathbf{C}$. $d(x, E)$ shall denote the distance from the point x to the set E, i.e.,

$$d(x,E) = \inf\{\|x-y\| \big| y \in E\},$$

while $d(A, B)$ shall denote the distance between the sets A and B, i.e.,

$$d(A,B) = \inf\left\{\|x-y\| \Big| x \in A, y \in B\right\} .$$

We shall have occasion to consider functions indexed by multi-indices, e.g. $f_0(x) = f_{0,\ldots,0}(x)$, *or* $f_\alpha(x) = f_{\alpha_1,\ldots,\alpha_n}(x)$. We shall suppose A to be a closed set in $\mathbf{R}^n$, bounded or unbounded. Suppose $f(x)$ is defined in A, and let $m \geq 0$ be an integer. We shall say: $f(x) = f_0(x)$ *is of class* C^m *in* A *in terms of the functions* $f_k(x)$ (with $|k| \leq m$) if the functions $f_k(x)$ are defined in A for all k with $|k| \leq m$ and satisfy, with $x, x' \in A$:

$$f_k(x') = \sum_{|\ell| \leq m - |k|} \frac{f_{k+\ell}(x)}{\ell!}(x' - x)^\ell + R_k(x'; x) ,$$

meaning:

$$
\begin{aligned}
f_{k_1\dots k_n}(x') \;=\; & \sum_{|\ell|\leq m-|k|} \frac{f_{k_1+\ell_1,\dots,k_n+\ell_n}(x)}{\ell_1!\ell_1!\dots\ell_n!}\,(x'_1-x_1)^{\ell_1}\dots(x'_n-x_n)^{\ell_n} \\
& + \; R_k(x';x) \;,
\end{aligned}
\tag{1}
$$

for each $f_k(x)$, with $|k| \leq m$; here $R_k(x';x)$ is assumed to satisfy: $\forall x^\circ \in A$, $\forall \varepsilon > 0$, $\exists \delta > 0 \ni$ if $x, x' \in A$ with $\|x - x^\circ\| < \delta$, $\|x' - x^\circ\| < \delta$ then

$$
|\; R_k(x';x)\;| \leq \; \|x-y\|^{m-|k|}\cdot\varepsilon \;. \tag{2}
$$

Note that if $m = 0$, these conditions (1) and (2) mean that $f(x)$ is continuous on the set A, and also that these conditions are satisfied automatically at all isolated points of A, regardless of how the $f_k(x)$ are defined there.

It is clear that the $f_k(x)$ are continuous and hence bounded in a neighbourhood of each point in A. Thus if f is of class C^m in A in terms of the $f_k(x)$ with $|k| \leq m$, then f is of class $C^{m'}, m' < m$ in terms of the $f_k(x)$ (with $|k| \leq m'$). We shall say that any arbitrary function $f(x)$ is of class C^{-1} in A, and that $f(x)$ is of class C^∞ in A in terms of the $f_k(x)$ if the $f_k(x)$ are defined for all k and f is of class C^m in A in terms of the $f_k(x)(|k| \leq m)$ for each integer $m \geq 0$.

Suppose $f(x)$ is defined in a region R and is of class C^m in terms of the $f_k(x)$ ($|k| \leq m$). Let $x = (x_1, \dots, x_n), x' = (x_1, \dots, x_h + \Delta x_h, \dots, x_n)$; then if $|k| < m$, we find:

$$
f_{k_1\dots k_n}(x') \;=\; f_{k_1\dots k_n}(x) + f_{k_1,\dots,k_n+1,\dots,k_n}(x)\Delta x_h + R^{(h)}_{k_1\dots k_n}(x';x)
$$

where $R^{(h)}_{k_1\dots k_n}(x';x)/\Delta x_h \to 0$ as $\Delta x_h \to 0$, which shows that

$$
\frac{\partial}{\partial x_h} f_{k_1\dots k_n}(x) \;=\; f_{k_1,\dots,k_h+1}(x) \quad \text{for} \quad |k| < m \;,
$$

in R. Hence in this case $f(x)$ is of class C^m in the ordinary sense, and the $f_k(x)$ are the partial derivatives of $f(x)$. Also Taylor's theorem shows that the converse is true.

We shall need a few lemmas, before turning to the proof of the approximation theorem of Whitney in question.

Lemma 1.1. *Let $w(z)$ be a continuous function of one variable defined in an interval I containing z_0, let B be a closed set in I, and let w_0' be a fixed number. Suppose* $\forall \varepsilon > 0 \quad \exists \delta > 0 \ni$

(1) if $z \in B$ and $|z - z_0| < \delta$ then $\left|\frac{w(z)-w(z_0)}{z-z_0'} - w_0'\right| < \varepsilon$;

(2) if $|z - z_0| < \delta$ and $z \notin B$ then $w'(z)$ exists and $|w'(z) - w_0'| < \varepsilon$.

Then $w(z)$ has a derivative at z_0 and $w'(z_0) = w_0'$.

The proof of this lemma is omitted since it is elementary. We shall make use of the functions denoted below by $\psi_k(x'; x)$. If $x \in A$, and $x' \in E(m < \infty)$ then we define:

$$\psi_k(x'; x) \;=\; \sum_{|\ell| \le m-|k|} \frac{f_{k+\ell}}{\ell!}(x' - x)^\ell \quad (|k| \le m) \; . \tag{3}$$

Thus $\psi_k(x'; x)$ is the value at x' of the polynomial of degree $\le m - |k|$ which approximates $f_k(x)$ to the $(m - |k|)^{th}$ order at x. For fixed x, it is a polynomial in x' given by Taylor's formula in terms of its value and derivatives at x. From (1) and (4) we then see that

$$f_k(x') \;=\; \psi_k(x'; x) + R_k(x' - x) \quad \text{for} \quad |k| \le m \; . \tag{4}$$

The ℓ^{th} derivative of $\psi_k(x'; x)$ (as a function of x') at x', is $\psi_{k+\ell}(x'; x)$. If we then ex-

press $\psi_k(x''; x)$ in terms of its value and derivatives at x', we obtain

$$\begin{aligned}\psi_k(x''-x) &= \sum_\ell \frac{\psi_{k+\ell}(x'-x)}{\ell!}(x''-x')^\ell \\ &= \sum_\ell \frac{(x''-x)^\ell}{\ell!}\sum_j \frac{f_{k+\ell+j}(x)}{j!}(x'-x)^j \;.\end{aligned} \tag{5}$$

This identity shows that

$$\begin{aligned}\psi_k(x''; x') &= \sum_\ell \frac{f_{k+\ell}(x')}{\ell!}(x''-x')^\ell \\ &= \sum_\ell \frac{(x''-'x)^\ell}{\ell!}\left[\sum_j \frac{f_{k+\ell+j}(x)}{j!}(x'-x)^j \;+\; R_{k+\ell}(x';x)\right] \\ &= \psi_k(x'';x) + \sum_\ell \frac{R_{k+\ell}(x';x)}{\ell!}(x''-x')^\ell \;.\end{aligned} \tag{6}$$

Our next objective is to construct a suitable C^∞ partition of unity on the set $\mathcal{C}(A)$. This will be done through several steps.

We define a function, which shall be denoted by $\Theta(x)$. Let R be the region defined by: $|x_h| < 1$, $(h = 1, 2, \ldots, n)$, R' be R minus the origin; and ∂R the boundary of R. We define the functions θ, θ', Θ as follows:

$$\begin{aligned}\theta(x) &= 2(1-x_1^2)(1-x_2^2)\ldots(1-x_n^2) - 1, \quad x \in R' \\ \theta'(x) &= \frac{\theta(x)}{1-\theta^2(x)}\;; \qquad\qquad x \in R'\end{aligned}$$

$$\Theta(x) = \begin{cases} e^{\theta'(x)}, & x \in R' \\ 0, & x \in \mathcal{C}(R)\;. \end{cases}$$

Then we see that $\theta(x)$ has the following properties:

$$-1 < \theta(x) < +1; \quad \theta(x) \to +1 \text{ as } x \to 0; \quad \theta(x) \to -1 \text{ as } x \to \partial R;$$

hence

$$\theta'(x) \to +\infty \text{ as } x \to 0; \quad \text{and} \quad \theta'(x) \to -\infty \text{ as } x \to \partial R\;.$$

Therefore

$$\Theta(x) \to +\infty \quad \text{to infinite order as} \quad x \to 0 \;,$$

and

$$\Theta(x) \to 0 \quad \text{to infinite order as} \quad x \to \partial R \;.$$

Also $\Theta(x)$ is C^∞ for $x \neq 0$. If $\Theta'(x) = \frac{1}{\Theta(x)}$ in R' and $\Theta'(x) = 0$ for $x = 0$ then $\Theta'(x)$ is C^∞ for $x \in R$.

The next step in the construction of a C^∞ partition of unity on $\mathcal{C}(A)$ is to define a suitable subdivision of this set. We first divide the space $\mathbf{R}^n$ into n-cubes of side 1 (we shall only consider cubes with sides parallel to the co-ordinate ones). Let K_0 be the set of all these cubes whose distances from A are at least $6n^{1/2}$ (if any). In general, having constructed the cubes of K_{s-1}, we divide each cube which is now present but not in $\cup_{i=1}^{s-1} K_i$ into 2^n cubes of side $\frac{1}{2^s}$; let K_s be the set of all these cubes whose distances from A are at least $\frac{6n^{1/2}}{2^s}$ (if any).

The following facts concerning this subdivision of $E - A$ will be needed.

Lemma 1.2. *The distance from any cube C of K_s to A is $< \frac{18n^{1/2}}{2^s} (s \geq 1)$.*

Proof. For it lies in a cube C' of the previous subdivision not belonging to K_{s-1} and whose distance from A is therefore $< \frac{6n^{1/2}}{2^{s-1}} = \frac{12n^{1/2}}{2^s}$.

Lemma 1.3. *Any cube C of K_s is separated from any cube C' of K_{s+2} by at least four cubes of K_{s+1}.*

Proof. This is true, because the distance $d(C, A) \geq \frac{12\sqrt{n}}{2^{s+1}}$, the distance from any point of C' to $A < \frac{9\sqrt{n}}{2^{s+1}}$, and the diameter of any cube of K_{s+1} is $\frac{\sqrt{n}}{2^{s+1}}$, which means that any

cube C' of K_{s+2} is separated from any cube of K_s by definitely more than 3 cubes of K_{s+1}; this number of intervening cubes of K_{s+1} has to be a *whole number* of cubes and hence at least four.

Our objective now is to introduce the functions $\phi_\nu(x), \nu = 1, 2, \ldots$. Let $y^1, y^2 \ldots$ be the set of all the vertices of $\cup_{i\geq 0} K_i$ arranged in a sequence;

$r_\nu = d(y^\nu, A) =$ distance from y^ν to A;

x^ν a fixed point of $A \ni d(x^\nu, y^\nu) = r_\nu$;

b_ν a fixed point of the side of the largest cube of $\cup_{i\geq 0} K_i$ with y^ν as a vertex;

I_ν be the set $\{x \in \mathbf{R}_n \mid |x_h - y_h^\nu| \leq b_\nu, \nu = 1, 2, \ldots, n\}$ and $B_\nu = \partial I_\nu$.

Then we define

$$\pi_\nu(x) = \Theta\left(\frac{x_1 - y_1^\nu}{b_\nu}, \ldots, \frac{x_n - y_n^\nu}{b_\nu}\right) \quad \text{in} \quad E - y^\nu \;;$$

$$\pi'_\nu(x) = \Theta'\left(\frac{x_1 - y_1^\nu}{b_\nu}, \ldots, \frac{x_n - y_n^\nu}{b_\nu}\right) \quad \text{in} \quad I_\nu - B_\nu;$$

$$\phi_\nu(x) = \begin{cases} \frac{\pi_\nu(x)}{\sum\limits_\lambda \pi_\lambda(x)} & \text{in } E - A, x \neq y^1, y^2, \ldots; \\ 1 & x = y^\nu; \\ 0 & x - y^\mu,\ \mu \neq \nu. \end{cases}$$

Lemma 1.4. *Let $y^* \in E - A$, and set $\delta_* = d(y^*, A)$ (or $\delta_* = d(y^*, x^0)$ for a point $x^\circ \in A$). Suppose y^* lies in a cube $C \in K_s$, and suppose I_ν, with centre y^ν, has points in common with C. Let $d_\nu = d(y^\nu, A)$ (or $d_\nu = d(y^\nu, x_0)$). Then*

$$\frac{d_*}{2} \leq d_\nu < 2\delta_* \,.$$

Proof. Let C' be a largest cube with y^ν as a vertex, and suppose $y' \in K_t$. Then $t \geq s - 1$. The diameter of C' as $\frac{\sqrt{n}}{2^t}$. Therefore the distance from y^ν to any point of I_ν is at

most $\frac{\sqrt{n}}{2^i} \le \frac{2\sqrt{n}}{2^s}$. The diameter of C is $\frac{\sqrt{n}}{2^s}$, hence $d(y^\nu, y^*) \le \frac{3\sqrt{n}}{2^s}$. Since $\delta_* \ge \frac{6\sqrt{n}}{2^s}$, it follows that $d(y^*, y^\nu) \le \frac{1}{2}\delta_*$.

Let $z \in A \ni d(y^*, z) = \delta_*$. Then $d_\nu \le d(y^\nu, z) \le \delta_* + d(y^*, y^\nu) \le \delta_* + \frac{\delta_*}{2} = \frac{3\delta_*}{2} < 2\delta_*$.

Also, $d(y^*, z) \le d(y^\nu, z) + d(y^*, y^\nu)$; hence $\delta_* \le d(y^\nu, z) + \frac{1}{2}\delta_*$, i.e., $\frac{\delta_*}{2} \le d(y^\nu, z)$. This is true for any $z \in A$, hence $\frac{\delta_*}{2} \le d_\nu$. Thus the Lemma is proved.

Next, each function $\pi_\nu(x)$ is > 0 in $I_\nu - B_\nu - y^\nu$, and only in this set. It tends to ∞ and 0 to infinite order as $x \to y^\nu$ and $x \to B_\nu$, respectively. Each point $x \in E - A$ is in I_ν° for some ν, hence $\pi_\nu(x) > 0$ for some ν, i.e., $\sum \pi_\lambda(x) > 0$ in $E - A$.

The function $\phi_\nu(x) \neq 0$ in $I_\nu - B_\nu$ and only on this set. Further

$$\sum_\nu \phi_\nu(x) = 1 \quad \text{if} \quad x \in E - A\,.$$

It is clear that $\phi_\nu(x)$ is C^∞ at points $x \neq y^\nu$. Let U_λ be a small neighbourhood of $y^\lambda, \lambda \neq \nu$. The function $\pi'_\lambda(x)$ is C^∞ in U^λ, hence so is $\phi_\nu = \frac{\pi'_\lambda \pi_\nu}{1+\pi'_\lambda \sum_\mu \pi_\mu}$ in U_λ. Similarly $\phi_\nu = \frac{1}{1+\pi'_\nu \sum_{\mu\neq\nu} \pi_\mu}$ is C^∞ in a small neighbourhood U_ν of y^ν. Thus $\phi_\nu(x)$ is C^∞ on $E - A$.

We shall next derive convenient estimates for the derivatives of the functions $\phi_\nu(x)$. Let C, C' be two closed cubes of $\cup_{i\ge 0} K_i$. The cubes C, C' are said *to be of the same type* if the sets in J' can be brought into coincidence with the sets in J by a translation and by stretching of the axes. There are at most a finite number, say d, of possible types of cubes, and for some number c, $\exists$ at most c sets I_ν with points in a given cube C.

Let C be a fixed cube of K_0, and $k \ge 0$ a fixed multi-index. Each $\phi_\nu(x)$ is C^∞, and $\phi_\nu(x) \neq 0$ only for a finite number of ν, hence

$$|D_k \phi_\nu(x)| < N_k(C) \;\; \forall x \in C \quad \text{and} \quad \forall \nu = 1, 2, \ldots$$

for some positive number $N_k(C)$.

Now let $C \in K_s$, and let $C \in K_0$ of the same type as C'. If $I_{\lambda'_1}, \ldots, I_{\lambda'_t}$ are the sets I_λ with points in C', let $I_{\lambda_1}, \ldots, I_{\lambda_t}$ be the corresponding sets with points in C which can be carried into the former by translation of the axes and stretching by a factor $\frac{1}{2^s}$. Each function ϕ_{λ_q} corresponding to I_{λ_q} thereby is mapped into the function

$$\phi_{\lambda'_q}(x) = \phi_{\lambda_q}\left\{y^{\lambda_q} + 2^s(x - y^{\lambda' q})\right\}$$

corresponding to $I_{\lambda'_q}$. On differentiating $|k|$ times w.r.t. x, we find

$$D_k \phi_{\lambda'_q}(x) = 2^{s|k|} D_k \phi_{\lambda_q}\big(y^{\lambda_q} = 2^s(x - y^{\lambda'_q})\big) \quad \forall x \in C' ,$$

and therefore, as $\phi_\nu(x) = 0$ in C' for $\nu \neq \lambda'_1, \ldots, \lambda'_t$,

$$\Big|D_k \phi_\nu(x)\Big| < 2^{s|k|} N_k(C) \quad \text{in} \quad C', \quad \nu = 1, 2, \ldots .$$

For a fixed k, there are at most d distinct values of $N_k(C)$; let N_k be the largest of these. The following lemma has thus been proved:

Lemma 1.5. *For any multi-index k, $\exists$ a number $N_k \ni$ if C is any cube of K_s then*

$$\Big|D_k \phi_\nu(x)\Big| < 2^{s|k|} N_k \quad \forall x \in C \quad \forall \nu = 1, 2, \ldots .$$

The next result establishes a $C^\infty-$ smooth extension of $f(x)$.

Lemma 1.6. *Let A be a closed set in $\mathbf{R}^n$ and let $f(x) = f_0(x)$ be of class C^m (m finite or infinite) in A in terms of the $f_k(x)$ ($|k| \le m$). Then $\exists$ function $g(x)$ which is C^∞ in $\mathbf{R}^n - A$, and has the properties:*

(1) $g(x) = f(x)$ in A,

(2) $D_k g(x) = f_k(x)$ in A, $|k| \le m$.

Proof. **Case I.** First suppose m finite. Let

$$g(x) = \begin{cases} \sum_\nu \phi_\nu(x)\psi(x;x^\nu), & x \in \mathbf{R}^n - A, \\ f(x), & x \in A, \end{cases}$$

where the $\phi_\nu(x)$ and $\psi(x, c^\nu) = \psi_0(x; x^\nu)$ have been defined above ($\nu = 1, 2, \ldots$). The $\phi_\nu(x)$ and $\psi(x; x^\nu)$ are C^∞ in $\mathbf{R}^n - A$, hence so is $g(x)$. The function $g(x) = f(x)$ is C^m at all inner points of A. We only have to show that $D_k g(x)$ exists, equals $f_k(x)$, and is continuous, at all boundary points of A, for $|k| \leq m$.

Let x° be a fixed boundary point of A, and let $\varepsilon \ni 0 < \varepsilon < 1$. Let $\eta > 0 \ni$

$$\eta < \min\left[\frac{\varepsilon}{6}, \frac{\varepsilon}{2c\{(m+2)!\}^n(108\sqrt{n})^m N}\right]$$

where $N = \max\{N_k, |k| \leq m\}$. Let $M > 0 \ni$

$$M > |f_k(x)|, \quad |k| \leq m, x \in A, \|x - x^\circ\| \leq 1,$$

and let

$$\delta < \min\left\{1, \frac{\varepsilon}{6(m+1)^n M}\right\}$$

so small that

$$\left|R_k(x; x^\circ)\right| \leq \|x - x^\circ\|^{m-|k|} \cdot \eta \; .$$

Now let $y^* \in \mathbf{R}^n - A \ni y^* \in B_{\delta/4}(x^\circ)$. We now assert:

Contention.

$$\left|D_k g(y^*) - f_k(x^\circ\right| < \varepsilon \quad \text{for } |k| \leq m \; .$$

To prove this contention, suppose $d(y^*, A) = \frac{\delta_*}{4}$, (so $\delta_* < \delta$), and let $x^* \in A \ni \|x^* - y^*\| = \frac{\delta_*}{4}$. Consider the function

$$\sum_{|\ell| \leq m-|k|} \frac{f_{k+\ell}(x)}{\ell!}(x^* - x^\circ)^\ell, \quad (|k| \leq m),$$

(which represents $\psi_k(x^*, x^\circ)$ in the notation adopted earlier). Each ℓ_h is $\leq m$, hence the above sum contains at most $(m+1)^n$ terms. On removing the term with $\ell_1 = \ldots = \ell_n = 0$ to the other side, we find in each remaining term a factor $(x_h - x_h^\circ)^{\ell_h}$ with $\ell_h > 0$. Each $|x_h^* - x_h^\circ| < \delta < 1$; hence

$$\left|\psi_k(x^*, x^\circ) - f_k(x^\circ)\right| < (m+1)^n M\delta < \frac{\varepsilon}{6} .$$

Also $\left|R_k(x^*, x^\circ)\right| < \eta < \frac{\varepsilon}{6}$, hence (using: $f_k(x^*) = \psi_k(x^*, x^\circ) + R_k(x^*, x^\circ)$)

$$\left|f_k(x^*) - f_k(x^\circ)\right| < \frac{\varepsilon}{3} .$$

Similarly, we find: $\left|\psi_k(y^*, x^*) - f_k(x^*)\right| < \frac{\varepsilon}{6}$. Hence

$$\left|\psi_k(y^*, x^*) - f_k(x^\circ)\right| < \frac{\varepsilon}{2} \qquad (|k| \leq m) .$$

Now suppose $y^* \in C \in K_s$; let $I_{\lambda_1}, \ldots, I_{\lambda_t}$ be those sets I_λ with points in C. Each corresponding point y^{λ_q} is at a distance $< \frac{\delta}{2}$ from x° (cf. Lemma 1.4), hence each corresponding point x^{λ_q} is at a distance $< \delta$ from x°. The same is true of x^*; hence using the characteristic property of $R_k(x^*, x^\circ)$, we find

$$\left|R_k(x^\nu; x^*)\right| \leq \|x^\nu - x^*\|^{m-|k|}\eta \qquad (\nu = \lambda_1, \ldots, \lambda_t) .$$

Now set

$$\zeta_{\nu,k}(x) \;=\; \psi_k(x; x^\nu) - \psi_k(x; x^*) \qquad (\nu = \lambda_1, \ldots, \lambda_t) ;$$

then as $\|x^\nu - x^*\| < \delta_*$, and $|x_h - x_h^\nu| < \delta_*$ for $x \in C$, it follows that $|(x - x^\nu)^\ell| < \delta_*^{|\ell|}$; hence (using the earlier identities for the function ψ_k):

$$\left|\zeta_{\nu,k}(x)\right| \;>\; (m+1)^n \delta_*^{m-|k|}\eta \qquad \forall x \in C \quad (\nu = \lambda_1, \ldots, \lambda_t) .$$

Since $\sum_{\nu} \phi_\nu(x) = 1$ in $\mathbf{R}^n - A$, we see that

$$g(x) \;=\; \psi(x;x^*) + \sum_{s=1}^{t} \phi_{\lambda_s}(x)\zeta_{\lambda_s;0}(x), \quad x \in C\ .$$

Now $D_k\psi(x;x^*) = \psi_k(x;x^\nu)$, hence $D_k\zeta_{\nu;0}(x) = \zeta_{\nu,k}(x)$. Therefore

$$D_k g(x) \;=\; \psi_k(x;x^*) + \sum_{s=1}^{t} \sum_{\ell} \binom{k}{\ell} D_\ell \phi_{\lambda_s}(x)\zeta_{\lambda_s;k-\ell}(x) \quad \text{in} \quad C\ .$$

Hence as $t \le c$, and $\binom{k_\ell}{\ell_h} \le m!$, we find

$$\Big| D_k g(x) - \psi_k(x;x^*)\ \Big| \;<\; \sum_{\ell} c\big((m+1)!\big)^n 2^{s|\ell|} N \delta_*^{m-(|k|+|\ell|)}\ .$$

To complete the proof we now use Lemma 1.2. Use the inequality established above: $|D_k g(y^*) - f_k(x^\circ)| < \varepsilon(|k| \le m)$, with $k = 0$; this shows that $g(x)$ is continuous throughout $\mathbf{R}^n$. Next let $k = (k_1, \ldots, k_n)$ with $|k| \le m$, and $k' = (k_1, \ldots, k_h + 1, \ldots, k_n)$.

Suppose $D_k g(x)$ is continuous in $\mathbf{R}^n$; we shall show that $D_{k'} g(x)$ exists and is continuous in $\mathbf{R}^n$. Let $x^\circ \;=\; (x_1^\circ,\ \ldots,\ x_n^\circ)$ be any boundary point, and write $z_0 \;=\; x_h^\circ$, $w(z) = w(x_h) \;=\; D_k g(x_1^\circ,\ \ldots,\ x_h,\ \ldots,\ x_n^\circ)$, $w_0' \;=\; f_{k'}(x^\circ)$. Let A^* be the set points of A with $x_p = x_p^\circ$ (for $p \ne h$). Then letting $\Delta x_h = x_h - x_h^\circ$, $x' = (x_1^\circ,\ \ldots,\ x_h^\circ + \Delta x_h,\ \ldots,\ x_n^\circ)$, and

$$f_{k_1,\ldots,k_n}(x^\circ + \Delta x_h) \;=\; f_{k_1,\ldots,k_n}(x^\circ) + f_{k_1,\ldots,k_h+1,\ldots,k_n}(x^\circ)\Delta x_h + R^{(h)}_{k_1\ldots k_n}(x';x^\circ)$$

and

$$\Big| D_{k'} g(y^*) - f_{k'}(x^\circ) \Big| < \varepsilon \qquad (\text{for} \quad |k| \le m)$$

show that the conditions of Lemma 1.2 are fulfilled; hence $\frac{\partial w(z_0)}{\partial x_h} = D_{k'} g(x^\circ)$ exists and equals $f_{k'}(x^\circ)$. Again the last inequality shows that $D_{k'} g(x)$ is continuous at x°. Hence $g(x)$ is C^m in $\mathbf{R}^n$.

Case II. Next we shall consider the case $m = \infty$.

For any given m, let $\psi_{m,k}(x';x)(|k| \leq m)$ be the function defined earlier:

$$\psi_{m,k}(x';x) \;=\; \sum_{|\ell| \leq m-|k|} \frac{f_{k+\ell}(x)}{\ell!}(x'-x)^{\ell} \qquad (|k| \leq m) \; .$$

Choose the axes so that the point $0 = (0, \ldots, 0) \in A$. Let S_p be the set of points of $\mathbf{R}^n$ whose distances from 0 are $\leq 2^p, p = 1, 2, \ldots$. Let $M_p = \max\{|f_k(x)| \;\big|\; |k| \leq p, x \in A \cap S_p\}$, and $N^{(p)} = \max\{N_k \;\big|\; |k| \leq p\}$. For each position integer p let $\delta_p \ni$

$$\delta_p < 1/\{2^{2p+1}c[(p+2)!]^n(36\sqrt{n})^p N^{(p)} M_{p+1}\}, \quad \delta_p < \frac{\delta_{p-1}}{2} \; .$$

The required extension $g^{\infty}(x)$ of $f(x)$ is determined as follows. For any given ν, determine $\gamma_\nu \ni \delta_{\gamma_\nu+1} \leq r_\nu < \delta_\nu$ (recall: $r_\nu = d(y^\nu, A)$). Let $\gamma_\nu = 0$ if $r_\nu > \delta_1$. Then we define

$$g^{\infty}(x) \;=\; \begin{cases} \sum_\nu \phi_\nu \psi_{\gamma_\nu,0}(x;x^\nu), & x \in \mathbf{R}^n - A, \\ f(x) \quad , & x \in A. \end{cases}$$

We shall establish an inequality for $D_k g^{\infty}(x)$ similar to one for $D_k g(x)$ (case I), for any k. Let $g^{(m)}(x)$ be the extension of class C^m obtained in the proof of case I ($m = 1, 2, \ldots$). Let x° be a boundary pont of A, and let $\varepsilon > 0$. Then choose $p \geq |k| + 2 \ni x^\circ \in S_p$ and $\ni \frac{1}{2^p} < \varepsilon$. Then choose $\delta < \delta_p \ni$ for any $y^* \in \mathbf{R}^n - A$ with $\|y^* - x^\circ\| < \delta$, we have

$$\left|D_k g^{(|k|)}(y^*) - f_k(x^\circ)\right| < \varepsilon \; .$$

We shall show:

$$\left|D_k g^{\infty}(y^*) - D_k g^{(|k|)}(y^*)\right| < \varepsilon \; .$$

Let $q \ni \delta_{q+1} \leq \delta_* < \delta_q$, where $\delta_* = d(y^*; A)$; then $q \geq p$. Define $C, K_s, I_{\lambda_1}, \ldots, I_{\lambda_t}$ as with preceding arguments (Lemma 1.5). Now for $\nu =$ any λ_h, $\delta_{\gamma_\nu+1} \leq r_\nu < 2\delta < 2\delta_p <$

δ_{p-1}, hence $\gamma_\nu + 1 > p - 1$, thus $\gamma_\nu > p - 2 \geq |k|$. We now write

$$\xi_\nu(x) \;=\; \psi_{\gamma_\nu;0}(x; x^\nu) - \psi_{|k|;0}(x; x^\nu) \qquad (\nu = \lambda_1, \ldots, \lambda_t)\,.$$

Using the definitions of $g^{(}|k|)(x)$ (from case I) and of $g^\infty(x)$ we find:

$$g^\infty(x) \;=\; g^{(|k|)}(x) + \sum_{u=1}^{t} \phi_{\lambda_u}(x)\xi_{\lambda_u}(x), \qquad x \in C\,.$$

Note that $D_j\xi_\nu(x) = \psi_{\gamma_\nu;j}(x; x^\nu) - \psi_{|k|;j}(x; x^\nu)$. Replacing k by j in the expression for $\psi_k(x'; x)$, we see that only those terms with $|\ell| \leq m - |j|$ occur. Now replace m by γ_ν and $|k|$ successively; then on subtracting we obtain

$$D_j\xi_\nu(x) \;=\; \sum_{|\ell|=|k|-|j|+1}^{\gamma_\nu - |j|} \frac{f_{j+\ell}(x^\nu)}{\ell!}(x - x^\nu)^\ell, \quad x \in C\,.$$

Also $r_\nu > \frac{\delta_*}{2}$, hence $r_\nu > \delta_{q+2}$, hence $\gamma_\nu \leq q + 1$ ($\nu = \lambda_1, \ldots, \lambda_t$). Therefore the number of terms in the sum is $< (q + 2)^n$, and in each term $|j| + |\ell| \leq q + 1$. Hence $|f_{j+\ell}(x^\nu)| \leq M_{q+1}$ in each term. Further $|x_h - x_h^\nu| < 2\delta_* < 2\delta_q$, and $|\ell| \geq |k| - |j| + 1$ in each term. Therefore

$$|D_j\xi_\nu(x)| < (q + 2)^n M_{q+1} 2^{q+1} \delta_*^{|k|-|j|} \delta_q, \qquad \forall x \in C\,.$$

Hence

$$|D_j g^{(\infty)}(x) - D_k g^{(|k|)}(x)| \leq \sum_{u=1}^{t} \sum_j \binom{k}{j} |D_{k-j}\phi_{\lambda_u}(x)|\; |D_j\zeta_{\lambda_u}(x)|$$

$$< c \sum_j \binom{k}{j} 2^{s(|k|-|j|)} N^{(|k|)} (q + 2)^n M_{q+1} 2^{q+1} \delta_*^{|k|-|j|} \delta_q \qquad \forall x \in C\,.$$

The distance $d(C, A) > \frac{\delta_*}{2}$, and is $< \frac{18\sqrt{n}}{2^s}$; hence $2^s < \frac{36\sqrt{n}}{\delta^*}$. Also $|k| < p \leq q$, therefore

$$|D_k g^{(\infty)}(x) - D_k g^{(|k|)}(x)| < c\big[(q + 2)!\big]^n (36\sqrt{n})^q N^{(q)} M_{q+1} 2^{q+1} \delta_q$$

$$< \frac{1}{2^q} < \varepsilon \qquad \forall x \in C\,,$$

in particular at y^*. Thus we find

$$|D_k g^{(\infty)}(y^*) - f_k(x^\circ)| < 2\varepsilon$$

for any point $y^* \in \mathbf{R}^n - A$ within distance δ of x°. Now again apply Lemma 1; and it follows that $D_k g^{(\infty)}(x)$ exists and is continuous throughout $\mathbf{R}^n$. This is true for every $k = (k_1, \ldots, k_n)$. This proves the Lemma.

We shall now turn to the following extension of the familiar Weierstrass approximation theorem.

Lemma 1.7. *Suppose $g(x)$ is of class C^m in $\mathbf{R}^n$ ($m < \infty$), and S be a compact set in $\mathbf{R}^n$. Then for each $\varepsilon > 0$ $\exists G$ analytic in $\mathbf{R}^n$ and satisfying:*

$$|D_k G(x) - D_k g(x)| < \varepsilon \qquad \forall x \in S, \quad |k| \leq m \; .$$

Proof. Let $R_b = \overline{B_k(0)}$ ($b \geq 0$), and consider the n-fold integral

$$\Phi(b) \;=\; T \int_{R_b} e^{-\|y\|^2} dy \;=\; T \int \cdots \int e^{-(y_1^2+y_2^2+\cdots+y_n^2)} dy_1 \, dy_2 \ldots dy_n$$

where T is $\ni \Phi(\infty = 1$. Then $0 \leq \Phi(b) \leq 1$ $\forall b$. Now replace y by κy and b by κb, and we obtain

$$\Phi(\kappa b) = T\kappa^n \int_{R_b} e^{-\kappa^2 \|y\|^2} dy \; .$$

Let $v(x)$ of class C^∞ be $\equiv$ 1 in S_1 $\equiv$ 0 outside some neighbourhood of S and $\ni$ $D_k v(x) \equiv 0$ in S $\forall k$ (cf. Lemma 1.6). Put $g'(x) = v(x)g(x)$, and let

$$G(x) \;=\; T\kappa^n \int_{\mathbf{R}^n} g'(y) e^{-\kappa^2 \|x-y\|^2} dy$$

where κ will be suitably chosen presently. Then $G(x)$ is analytic in $\mathbf{R}^n$. The function $\|x-y\|^2$ is a function of $x-y$ only and differentiating under the integral sign gives

$$\begin{aligned} D_k g(x) &= T\kappa^n \int_{\mathbf{R}^n} g'(y) D_k^{(x)} e^{-\kappa^2\|x-y\|^2} dy \\ &= (-1)^{|k|} T\kappa^n \int_{\mathbf{R}^n} g'(y) D_k^{(y)} e^{-\kappa^2\|x-y\|^2} dy \ , \end{aligned}$$

where $D_k^{(x)}, D_k^{(y)}$ denote differentiation with respect to x and y, respectively. Then integration by parts $|k|$ times yields

$$D_k g(x) \ = \ T\kappa^n \int_{\mathbf{R}^n} D_k g'(y) e^{-\kappa^2\|x-y\|^2} dy \ .$$

If we recall that $\Phi(\infty) = 1$, we see that

$$D_k g(x) - D_k g'(x) \ = \ T\kappa^n \int_{\mathbf{R}^n} \{D_k g'(y) - D_k g'(x)\} e^{-\kappa^2\|x-y\|^2} dy \ .$$

Now let $M > 0$ so large that

$$|D_k g'(x)| < M \qquad \forall x \in \mathbf{R}^n \qquad (|k| \le m) \ .$$

The functions $D_k g'(x)$ are uniformly continuous on $\mathbf{R}^n$, hence $\exists\ \delta > 0 \ni$

$$\left| D_k g'(y) - D_k g'(x) \right| < \frac{\varepsilon}{2} \quad \forall y \in B_r(x), \quad |k| \le m \ .$$

Then let $\kappa > 0$ so large that

$$1 - \Phi(\kappa\delta) < \frac{\varepsilon}{4M} \ .$$

Let $U = B_\delta(x)$, and denote by J_1 and J_2 the regions formed by replacing the domain of integration on the right side of $|D_k G(x) - D_k g'(x)|$ above, by U, and $\mathbf{R}^n - U$, respectively; we then obtain:

$$\begin{aligned} |J_1| &< T\kappa^n \int_U \frac{\varepsilon}{2} e^{-\kappa^2\|x-y\|^2} dy = \frac{\varepsilon}{2}\Phi(\kappa\delta) < \frac{\varepsilon}{2}, \\ |J_2| &< T\kappa^n \int_{\mathbf{R}^n - U} 2M e^{-\kappa^2\|x-y\|^2} dy = 2M\{1 - \Phi(\kappa\delta)\} < \frac{\varepsilon}{2}, \end{aligned}$$

hence

$$\left|D_k G(x) - D_k g'(x)\right| < \varepsilon \qquad \forall x \in S, \quad |k| \le m \ .$$

This completes the proof of Lemma 1.7.

The preceding Lemma can be generalised to yield the next theorem which is the main theorem of Whitney that we are aiming at in this Chapter.

Theorem 1.8. *Let R be an open set in $\mathbf{R}^n$, and $R_1, R_2, \ldots$ bounded open sets $\ni \bigcup_{n\ge1} R_n = R$, and $\ni \overline{R}_p \subset P_{p+1}$ for each p. Suppose g is defined and of class C^m in R (m finite or finite), and suppose $\varepsilon_1 \ge \varepsilon_2 \ge \ldots$ are given positive number. Then $\exists$ analytic function $G(x)$ in R, satisfying*

$$\left|D_k G(x) - D_k g(x)\right| < \varepsilon_p \quad \forall x \in R - R_p, \quad |k| \le \alpha_p, \quad p = 1, 2, \ldots,$$

where

$$\alpha_p = \begin{cases} m & \text{if } m \text{ is finite} \\ p & \text{if } m = \infty, \ p = 1, 2, \ldots \ . \end{cases}$$

Note that if $R_1, \ldots, R_q$ are empty, then this statement means

$$\left|D_k G(x) - D_k g(x)\right| < \varepsilon_q \quad \text{in} \quad R, \quad |k| \le \alpha_q \ .$$

Proof. Consider the closed set

$$\begin{aligned} &\overline{R}_{p-1} \cup (\overline{R}_{p+1} - R_p) \cup (\mathbf{R}^n - R_{p+2}) \\ &= Q'_p \cup Q_p \cup Q''_p \ . \end{aligned}$$

In Lemma 1.6 we replace the closed set A by this set, and replace $f(x)$ (of Lemma 1.6) by a function which is $\equiv 1$ in Q_p, and $\equiv 0$ in $Q'_p \cup Q''_p$. Then for each p, let $u_p(x)$ be a function, C^∞ in $\mathbf{R}^n$, $\ni$

$$u_p(x) = \begin{cases} 1 & \text{in } Q_p, \\ 0 & \text{in } Q'_p \cup Q''_p \end{cases}$$

and $D_k u_p(x) = 0$ in $Q'_p \cup Q_p \cup Q''_p$ $(|k| > 0)$. (If $R_{p+1} = \emptyset$ we put $u_p(x) \equiv 0$; if $R_{p+1} \neq \emptyset$ but $R_{p-1} = \emptyset$, then $u_p(x) = 0$ in Q''_p, and $= 1$ in $\overline{R}_{p+1}$.) Now let $Z_p \geq 1$ be a number $\ni$

$$\left|D_k u_p(x)\right| < Z_p \qquad \forall x \in \mathbf{R}^n, \quad |k| \leq \alpha_p, \quad p = 1, 2, \ldots .$$

Now define, successively, the analytic functions $G_1(x), G_2(x), \ldots,$ by:

$$G_p(x) = T\kappa_p^n \int_{\mathbf{R}^n} u_p(y)\big[g(y) - \{G_1(y) + \cdots + G_{p-1}(y)\}\big] e^{-\kappa_p^2 \|x-y\|^2} dy .$$

(If $p = 1$, the factor in brackets is $g(y)$.) The constant κ_p is so chosen that if we set

$$H_p(x) = u_p(x)\big[g(y) - \{G_1(x) + \cdots + G_{p-1}(x)\}\big]$$

then

$$\left|D_k G_p(x) - D_k H_p(x)\right| < \beta'_p = \varepsilon_{p+1} / \Big\{2^{p+2}\big[(\alpha_{p+1} + 1)!\big]^n Z_{p+1}\Big\}$$

in $\overline{R}_{p+1}(|k| \leq \alpha_{p+1}$; cf. Lemma 7; the constant κ_p will be further restricted a little later on). From the definition of $u_p(x)$ we see from the last two relations that

$$\left|D_k g(x) - D_k\{G_1(x) + \cdots + G_p(x)\}\right| < \beta'_p < \frac{\varepsilon_p}{2} \quad \text{in} \quad Q_p(|k| \leq \alpha_{p+1}) .$$

We then differentiate $H_p(x)$, replacing p by $p-1$ in the preceding relations and we see that (cf. proof of Lemma 1.6, $m < \infty$):

$$\left|D_k H_p(x)\right| < \big[(\alpha_p + 1)!\big]^n Z_p \beta'_{p-1} = \frac{\varepsilon_p}{2^{p+1}} \quad \text{in} \quad Q_{p-1} \quad (|k| \leq \alpha_p) .$$

The function $u_p(x)$ and its derivatives are 0 in R_{p-1}, hence the preceding holds also in R_{p-1}. Hence

$$\left|D_k G_p(x)\right| < \frac{\varepsilon_p}{2^p} \quad \text{in} \quad \overline{R}_p \quad (|k| \leq \alpha_p) .$$

Now let

$$G(x) = G_1(x) + G_2(x) + \cdots .$$

We shall show that $G(x)$ yields the required approximation to $g(x)$. We note that $D_k\{G_1(x) + \cdots + G_p(x)\}$ converges uniformly in any compact subset of R $(|k| \leq m)$; hence $G(x)$ is defined in R and

$$D_k G(x) = \sum_{i=1}^{\infty} D_k G_i(x) \qquad \forall x \in R \qquad (|k| \leq m) .$$

Next since $|D_k G_p(x)| < \frac{\varepsilon_p}{2^p}$ in $\overline{R}_p(|k| \leq \alpha_p)$, it follows that

$$\left| D_k G_{p+1}(x) + D_k G_{p+2}(x) + \cdots \right| < \frac{\varepsilon_{p+1}}{2^{p+1}} + \frac{\varepsilon_{p+2}}{2^{p+2}} + \cdots$$

$$< \varepsilon_p\Big(\frac{1}{2^2} + \frac{1}{2^3} + \cdots\Big) = \frac{\varepsilon_p}{2} \qquad \forall x \quad \overline{R}_{p+1} \quad (|k| \leq \alpha_{p+1}) .$$

Hence from a previous estimate we obtain:

$$\left| D_k G(x) - D_k g(x) \right| < \varepsilon_p \qquad \forall x \in Q_p \quad (|k| \leq \alpha_p) .$$

This proves part of the Theorem.

Now we want to show that $G(x)$ is analytic in R. We extend the definition of each $G_p(x)$ to complex values of $x = (x_1' + i x_1'', \ldots, x_n' + i x_n'')$, using the definition of $G_p(x)$. Consider the analytic function of x:

$$r_{x,y}^2 = \sum_h (x_h - y_h)^2 = \sum_h \{(y_h' - x_h') + i(y_h'' - x_h'')\}^2 .$$

The domain of integration in the integral defining $G_p(x)$ is real, hence $y_h'' = 0$, hence

$$\mathrm{Re}(r_{x,y}^2) = \sum_{h=1}^{n} \{x_h' - y_h')^2 - x_h''^2\} .$$

Now let $x^\circ \in R$ and U be the *complex* open ball $B_\rho(x^\circ)$ where $\rho > 0$ is so small that the real points in the complex open ball $B_{3\rho}(x^\circ)$ lie in some R_q; we take q so that $3\rho^2 > \frac{1}{2^q}$. Now if $p > q$, $x \in U$ and $y \in R - R_{p-1}$, then $\sum x_h''^2 < \rho^2$ and $\sum (x'_h - y'_h)^2 \geq 4\rho^2$, hence

$$\mathrm{Re}\left(r_{xy}^2\right) > 3\rho^2 .$$

Furthermore $H_p(y) = 0$ in R_q and in $\mathbf{R}^n - R_{p+2}$ for $p > q$. Hence if $M'_p = \max |H_p(y)|$ (recall: $H_p(y)$ is determined before κ_p), and if $V_p =$ volume of $R_p (p = 1, 2, \ldots)$, then

$$\begin{aligned}\left|G_p(x' + ix'')\right| &< T\kappa_p^n \int_{R_{p+2} - R_{p-1}} M'_p e^{-3\kappa_p^2\rho^2}\, dy \\ &< T\kappa_p^n e^{-\kappa_p^2/2^b} M'_p V_{p+2}\end{aligned}$$

if $x \in U$ and $p > q$. Now choose κ_p successively for $p = 1, 2, \ldots$ so that the term on the right side in the preceding inequality is $< \frac{1}{2^p}$, then the series defining $G(x)$ converges uniformly in a complex neighbourhood of any point of R. Hence $G(x)$ is analytic in R. This completes the proof of the theorem.

§2. C^∞ approximation in a finite dimensional space

We shall now turn to the proof of a slightly weaker theorem than Whitney's on strong approximation in a finite dimensional space. The methods used, however, are sufficiently different from those used by Whitney to merit special attention. It is now necessary to define the concept of approximation in a strong (or fine) topology. The compact open topology (Chapter I) is suitable for investigating closeness of maps over a compact set E, whereas if E is *not* compact then the compact open topology does not yield sufficient control over the behaviour of a map, and a strong (or fine) topology is more useful. To be more precise, the word "fine" is prefixed by words such as "C^0" or "C^k", as will be made clear presently.

We shall define the concept of C^j-fine topology on $C^k(U, Y)$ where U is a nonempty open set in a Banach space X, Y is a Banach space, k is a given nonnegative integer and j is an integer $\ni 0 \leq j \leq k$. When speaking of differentiability of a map defined in a Banach space we shall always understand Fréchet differentiability unless specifically stated otherwise. As is customary E^j denotes the j-fold Cartesian product $\underbrace{E \times \cdots \times E}_{j}$. C^j-smooth means j-times differentiable and the jth derivative, as an element in the space $L^j(E, F)$ of continuous j-linear operators $E^j \to F$, depending on $x \in U$, is continuous in U hence is symmetric and belongs to the space $L^j_s(E, F)$, of continuous j linear symmetric operators $E^j \to F$. C^∞-smooth means C^j-smooth for every integer $j \geq 0$. $D^j\psi$ denotes the jth successive derivative of ψ if it exists. $C^k(U, Y)$ is the space of C^k-smooth mappings $U \to Y$. The *C^j-fine topology* on $C^k(U, Y)$, $0 \leq j \leq k$ is defined to be

the topology for which sets of the form:

$$N(h,\eta) = \Big[g \in C^k(U,Y) \;\Big|\; \forall \text{ integers } i \in [0,j], \|D^i g(x) - D^i h(x)\| < \eta(x) \quad \forall x \in U \Big]$$

where $h(\cdot) \in C^k(U,Y)$ and $\eta(\cdot)$ is an arbitrary positive continuous function on U, form a base. The C^k-fine topology on $C^k(U,Y)$ is also denoted by $C^k_s(U,Y)$.

We should point out that although the preceding definition is stated for Banach spaces we are emphasizing results in finite dimensional spaces in this chapter. So we shall now onward in this section deal with functions defined in an open set U in $\mathbf{R}^m$ with values in $\mathbf{R}^n$.

Before proceeding to the theorem in question and its proof we shall need some preliminary lemmas. Let

$$\alpha(t) = \begin{cases} e^{-\frac{1}{(t-a)(b-t)}}, & 0 < a < t < b, \\ 0 \;, & t \notin (a,b); \end{cases}$$

$$\beta(x) = \frac{\int_x^b \alpha(t)dt}{\int_a^b \alpha(t)dt}, \quad x \in \mathbf{R}^1;$$

and $\phi(x) = \beta(\|x\|^2)$. The function $\beta(\cdot)$ is C^∞ on $\mathbf{R}^1, \beta(\cdot) \equiv 1$ on $(-\infty, a), \beta(\cdot) \equiv 0$ on (b,∞) while $\phi(\cdot)$ is C^∞ on $\mathbf{R}^m, \phi(\cdot) \equiv 1$ on $\overline{B_{\sqrt{a}}(0)}$ and $\phi(\cdot) \equiv 0$ outside $B_{\sqrt{b}}(0)$. This function $\phi(\cdot)$ also enables us to construct a C^∞-function $\theta : \mathbf{R}^m \to \mathbf{R}^1 \ni \theta = 0$ outside a compact set and its Lebesgue integral in $\mathbf{R}^m$ equals 1.

By the *support* of a continuous real-valued function f, denoted by Supp f, is meant the set $\overline{f^{-1}(\mathbf{R}^1 - \{0\})}$, so the complement of this set, $\mathcal{C}(\text{Supp } f)$ is the largest open set on which $f = 0$. Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be an open cover of $\mathcal{U}$. By a *C^j-partition of unity subordinate* to $\mathcal{U}$ we mean a family of C^j-maps $\lambda_\alpha : U \to [0,1], \alpha \in A, \ni$

(i) supp $\lambda_\alpha \subset U_\alpha, \quad \alpha \in A$;

(ii) $\{\text{supp } \lambda_\alpha\}_{\alpha \in A}$ is a locally finite family; and

(iii) $\sum_{\alpha \in A} \lambda_\alpha(x) = 1, \quad x \in U.$

The local finiteness property of $\{\text{Supp } \lambda_\alpha\}_{\alpha \in A}$ ensures that each point of U has a neighbourhood on which all except a finite number of λ_α are 0, and the sum is locally a finite sum.

We note that condition (iii) ensures that

$$U = \bigcup_{\alpha \in A} (\text{supp } \lambda_\alpha)^0 ,$$

where E° denotes the interior of E. We also note the following simple observation:

If $\mathcal{V} = \{V_\alpha\}_{\alpha \in A}$ is an open cover of U which refines $\mathcal{U} = \{U_\beta\}_{\beta \in B}$, and if $\mathcal{V}$ has a subordinate C^∞-partition of unity, then so does $\mathcal{U}$.

Lemma 2.1. *Every open cover of U has a subordinate C^j-partition of unity, for any given $j \ni 0 \leq j \leq \infty$.*

This result is well-known; for a sketch of the proof see §3 in Chapter III.

To approximate C^r-maps by C^∞-maps in the strong topology we need to approximate locally f on the U_i, where $\{U_i\}_{i \in \Lambda}$ is an open covering, by C^∞ maps whose derivatives up to order r uniformly approximate those of f. In a finite dimensional space this is achieved by using the technique of convolutions, which we shall explain next.

Let $\theta : \mathbf{R}^m \to \mathbf{R}$ be a function with compact support. There is a smallest $\sigma \geq 0 \ni$ supp θ is contained in the closed ball $\overline{B_\sigma(0)} \subset \mathbf{R}^m$. We call σ the *support radius* of θ.

Now suppose $U \subset \mathbf{R}^m$ is an open set and $f : U \to \mathbf{R}^n$ a map. If $\theta : \mathbf{R}^m \to \mathbf{R}$ has compact support then we define the *convolution of f* by θ to be the function denoted by

$\theta * f : U_\sigma \to \mathbf{R}^n$, defined by

$$\theta * f(x) = \int_{B_\sigma(0)} \theta(y) f(x-y) dy, \qquad x \in U_\sigma ,$$

where $U_\sigma = \{x \in U \mid B_\sigma(x) \subset U\}$. Here we understand integration to be in the Lebesgue sense, dy denoting Lebesgue measure in $\mathbf{R}^m$. We note that the integrand is 0 on the boundary of $B_\sigma(0)$; we define the integrand to be 0 outside $B_\sigma(0)$ and thus extend it to all of $\mathbf{R}^m$. Then

$$\theta * f(x) = \int_{\mathbf{R}^m} \theta(y) f(x-y) dy, \qquad x \in U_\sigma .$$

Now for a fixed $x \in U_\sigma$ we make the change of variable: $z = x - y$. Then

$$\theta * f(x) = \int_{B_\sigma(x)} \theta(x-z) f(z) dz = \int_{\mathbf{R}^m} \theta(x-z) f(z) dz, \qquad x \in U_\sigma ,$$

where the integrand is defined to be 0 outside $B_\sigma(x)$.

A function $\theta : \mathbf{R}^m \to \mathbf{R}$ is called a *convolution kernel* if $\theta \geq 0$, θ has compact support and $\int_{\mathbf{R}^m} \theta = 1$. Earlier we noted that such kernels which are further C^∞, do exist. Since $\theta * f(x)$ can be considered to be a weighted average of values of f near x, it is plausible to expect that $\theta * f$ might be an approximation to f in a neighbourhood of x, and the approximation might be smooth.

We shall use the notation

$$\|f\|_{r,K} = \sup\left\{\|D^k f(x)\| \;\middle|\; x \in K, \quad 0 \leq k \leq r\right\}$$

if $f : U \to \mathbf{R}^n$ is $C^r, U \subset \mathbf{R}^m$ is open, $K \subset U$ is any subset, and $\|D^k f(x)\|$ is the norm of the kth derivative (as a continuous k-linear symmetric operator) of f at x. $\|D^\circ f(x)\|$ means $\|f(x)\|$.

We now need the following result.

Lemma 2.2. *Let $\theta : \mathbf{R}^m \to \mathbf{R}$ have support radius $\sigma > 0$. Let $U \subset \mathbf{R}^m$ be an open set and $f : U \to \mathbf{R}^n$ be a continuous function. Then $\theta * f : U_\sigma \to \mathbf{R}^n$ has the following properties:*

(a) *If $\theta|_{\mathrm{supp}(\theta)^\circ}$ is $C^k, 1 \leq k \leq \infty$, then so is $\theta * f$, and for each finite k*

$$D^k(\theta * f)(x) = D^k\theta * f(x) .$$

(b) *If f is C^k Then*

$$D^k(\theta * f) = \theta * D^k(f) .$$

(c) *Suppose $f \in C^r, 0 \leq r \leq \infty$. Let $K \subset U$ be compact. Given $\varepsilon > 0 \exists \sigma > 0 \ni$ $K \subset U_\sigma$ and θ is a C^r-convolution kernel with support radius σ, then $\theta * f \in C^r$ and*

$$\|\theta * f - f\|_{r,K} < \varepsilon .$$

Proof. The part (b) follows by differentiating under the integral sign. The part (a) follows by a change of variable, viz. $z = x - y$. (c) It suffices to let $r = 0$. Since dist $(K, \mathbf{R}^m - U) > 0$, we can choose $\sigma > 0$ sufficiently small so that $K \subset U_\sigma$. Also let σ sufficiently small so that, if $x \in K$ and $\|x - y\| \leq \sigma$ then $|f(x) - f(y)| < \varepsilon$.

Now use the fact that $\int_{\mathbf{R}^m} \theta = 1$. On integrating over $\mathbf{R}^m$, we obtain:

$$\begin{aligned} |\theta * f(x) - f(x)\Big| &= \Big|\int \theta(y)\{f(x-y) - f(x)\}dy\Big| \\ &\leq \int \theta(y)|f(x-y) - f(x)|dy \leq \varepsilon \int \theta(y)dy = \varepsilon . \end{aligned}$$

This proves the Lemma.

Thus we see that a C^r map from an open subset U of $\mathbf{R}^m$ to $\mathbf{R}^n$ can be C^r approximated by C^∞ maps in neighbourhoods of compact subsets of U. We shall follow Kurzweil (see[30]) for the proof of the next lemma.

Lemma 2.3. *Suppose X is a metric space and $G \subset X$ an open set. Suppose $G = \bigcup_{i=1}^{\infty} B_i$ where we have written B_i for $B_{\delta_i}(x_i)$ = the open ball in the centre x_i and radius $\delta_i, i = 1, 2, \ldots$. Then $\exists$ locally finite open covering $\{V_i\}_{i=1}^{\infty}$ of $G \ni V_i \subset B_i, i = 1, 2, \ldots$.*

Remark. The V_i in this lemma will turn out to have further properties: viz. each V_i will be "scalloped" to use the terminology of Lang [32] p. 35.

Proof. Choose a sequence of positive numbers $\{\varepsilon_i\}_{i=1}^{\infty} \ni 3\varepsilon_i < \delta_i$, $i = 1, 2, \ldots$, and $1 > \varepsilon_1 < \varepsilon_2 > \cdots$ with $\lim_{i \to \infty} \varepsilon_i = 0$. Then define $V_1 = B_1$, $V_2 = B_2 \cap \mathcal{C}\overline{B_{\delta_1 - \varepsilon_2}(x_1)}$; and having defined V_{j-1}, define

$$V_1 = B_j \cap \mathcal{C}\overline{B_{\varepsilon_1 - \varepsilon_j}(x_1)} \cap \mathcal{C}\overline{B_{\delta_2 - \varepsilon_j}(x_2)} \cap \ldots \cap \mathcal{C}\overline{B_{\varepsilon_{j-1} - \varepsilon_j}} \ .$$

Clearly $G = \bigcup_{j=\infty}^{\infty} V_j$, and $V_j \subset B_j, j = 1, 2, \ldots,$.

Let $y \in G$, and let $k - \inf[m | y \in B_m]$. There is an integer $\ell > k \ni y \in B_{\delta_k - 3\varepsilon_\ell}(x_k)$. Then

$$B_{\delta_k - 3\varepsilon_\ell}(x_k) \cap \mathcal{C}\overline{B_{\delta_k - \varepsilon_j}(x_k)} = \emptyset \qquad \text{for} \quad j \geq \ell \ .$$

Hence

$$B_{\delta_k - 3\varepsilon_\ell}(x_k) \cap V_j = \emptyset \qquad \text{for} \quad j \geq \ell \ .$$

This means that $\exists$ neighbourhood viz. $B_{\delta_k - 3\varepsilon_\ell}(x_k)$ of the point y which intersects only a finite number of the sets V_j, $j = 1, 2, \ldots$. This shows that the covering $\{V_j\}_{j=1}^{\infty}$ is locally finite. This proves the lemma.

Lemma 2.4. *Suppose $G \subset \mathbf{R}^m$ is an open set. Then $\exists$ countable locally finite covering $\{C_i\}_{i=1}^{\infty}$ of G by compact set C_i.*

Proof. The proof is a slight modification of the proof of the last lemma. For each $x \in G \exists$ open ball $B_{2\delta}(x) \subset G$. Then $\exists$ countable subcollection $\{B_{\delta_i}(x_i)\}_{i=1}^{\infty} \ni G = \bigcup_{i=1}^{\infty} B_i$, where we set $B_i = B_{\delta_i}(x_i)$. Then $G = \bigcup_{i=1}^{\infty} \overline{B_i}$, and each $\overline{B_i}$ is compact.

Let $\{\varepsilon_i\}_{i=1}^{\infty}$ be a sequence of positive numbers $\ni 3\varepsilon_i < \delta_i$, $1 > \varepsilon_1 > \varepsilon_2 > \cdots$, and $\lim_{i\to\infty} \varepsilon_i = 0$. Define

$$C_1 = \overline{B}_1, \quad C_2 = \overline{B}_2 \cap \mathcal{C}B_{\delta_1-\varepsilon_2}(x_1) \ ;$$

generally having defined C_{j-1}, define

$$C_j = \overline{B}_j \cap \mathcal{C}B_{\delta_1-\varepsilon_j}(x_1) \cap \mathcal{C}B_{\delta_2-\varepsilon_j}(x_2) \cap \ldots \cap \mathcal{C}B_{\delta_{j-1}-\varepsilon_j}(x_{j-1}) \, .$$

Then $G = \bigcup_{j=1}^{\infty} C_j$, $C_j \subset \overline{B}_j$ and the C_j are compact, for $j = 1, 2, \ldots$.

Let $y \in G$, and let $k = \inf\,[m | y \in \overline{B}_{\delta_m}]$. There is an integer $\ell > k \ni y \in B_{\delta_k - 3\epsilon_\ell}(x_k)$. Then

$$B_{\delta_k-3\epsilon_\ell}(x_k) \cap \mathcal{C}B_{\delta_k-\epsilon_j}(x_k) = \emptyset \qquad \text{for} \quad j \geq \ell \ ;$$

hence

$$B_{\delta_k-3\epsilon_\ell}(x_k) \cup C_j = \emptyset \qquad \text{for} \quad j \geq \ell \, .$$

This means that $\exists$ neighbourhood viz. $B_{\delta_k-3\epsilon_\ell}(x_k)$ of the point y which intersects only a finite number of the $C_j, j = 1, 2, \ldots$, i.e., the covering $\{C_j\}_{j=1}^{\infty}$ of G is locally finite. This proves the lemma.

Theorem 2.5. *Let $U \subset \mathbf{R}^m$ be an open set. Then $C^\infty(U, \mathbf{R}^n)$ is dense in $C^r(U, \mathbf{R}^n)$ for $0 \leq r < \infty$, in the C^r-fine topology.*

Proof. Let $f \in C(U, \mathbf{R}^n)$. Let $K = \{K_i\}_{i \in \wedge}$ be a locally finite family of compact sets covering U, let $e = \{\varepsilon_i\}_{i \in \wedge}$ be a family of positive numbers and define the sets $N(f, \mathbf{K}, e)$ to be the set of functions $g \in C^r(U, \mathbf{R}^n) \ni$

$$\forall i \in \wedge, \qquad \|g - f\|_{r,K_i} < \varepsilon_i \ .$$

We want to show that with a fixed family $\mathbf{K}$, if we let $\varepsilon(\cdot) > 0$ be an arbitrary positive continuous function on U, and $\varepsilon_i = \min \varepsilon(\cdot)$ on K_i where $K_i \in \mathbf{K}, i \in \wedge$, then

$$C^\infty(U, \mathbf{R}^n) \ \cap \ N(f, \mathbf{K}, e) \neq \emptyset \ .$$

Let $\{\lambda_i\}_{i \ni \wedge}$ be a C^∞-partition of unity on $U \ \ni$ $\text{supp}(\lambda_i)$ is compact and contains K_i.

Given any set of positive numbers $\{\alpha_i\}_{i \in \wedge} \exists$ C^∞ maps $g_i : U_i \to \mathbf{R}^n$ (defined on suitable open sets $U_i \supset K_i, i \in \wedge$) satisfying:

$$\|g_i - f\|_{r,K_i} < \alpha_i, \qquad i \in \wedge \ .$$

Then we define $g : U \to \mathbf{R}$ by: $g(x) = \sum_i \lambda_i(x) g_i(x)$. Then clearly $g \in C^\infty$. We need to estimate $\|D^k g(x) - D^k f(x)\|$, for $0 \leq k \leq r$.

We note the following generalised Leibnitz formula: if $\lambda : U \to \mathbf{R}$, and $h : U \to \mathbf{R}^n$ are C^k functions and $\psi(x) = \lambda(x) h(x)$, then

$$D^k(\lambda h)(x) \ = \ D^k \lambda(x) h(x) + \binom{k}{1} D^{k-1} \lambda(x) D h(x) + \cdots + \lambda(x) D^k h(x)$$

where the products are tensor products. Hence

$$\|D^k(\lambda h)(x)\| \leq A_k \max_{0 \leq p \leq k} \|D^p \lambda(x)\| \max_{0 \leq q \leq k} \|D^q h(x)\|$$

for a suitable positive constant A_k-independent of x, λ, and h. Let $A = \max_{0 \le k \le r}\{A_k\}$, and for fixed $i \in \wedge$, $\wedge_i = \{j \in \wedge \mid K_i \cup K_j \neq \emptyset\}$. This is a finite set, with cardinality, m_i, say. Put

$$\begin{aligned} \mu_i &= \max\{\|\lambda_j\|\}_{r,K_i} \mid j \in \wedge_i\} \\ \beta_i &= \max\{\alpha_j \mid j \in \wedge_i\}\,. \end{aligned}$$

Then for $x \in K_i$, j varying over $\wedge_i$, and $0 \le k \le r$, we have

$$\begin{aligned} \|D^k g(x) - D^k f(x)\| &= \|\sum_j D^k(\lambda_j g_j - \lambda_j f)(x)\| \\ &\le \sum_j \|D^k\{\lambda_j(g_j - f)\}(x)\| \le m_i A \mu_i \beta_i\,. \end{aligned}$$

Now we choose the numbers α_i so that $m_i A \mu_i \beta_i < \varepsilon_i$. With this choice of α_i, $i \in \wedge$, we find: $\|g - f\|_{r,K_i} < \varepsilon_i$. This proves Theorem 2.5.

We shall now turn to another aspect of strong approximation - strong approximation by C^∞ maps with some further special properties, viz. transversality. Such results are more natural in the context of differentiable manifolds Thus some of the results are formulated for manifolds. One additional technique has now to be used, viz. the Morse-Sard theorem. Some definition and preliminary results are necessary.

An *n-cube* $C \subset \mathbf{R}^n$ of edge $\lambda > 0$ is a product

$$C = I_1 \times \cdots \times I_\lambda \subset \mathbf{R}^n$$

of closed intervals each of length λ: say $I_j = [a_j, a_j + \lambda] \subset \mathbf{R}$. The *(n)-measure* of C is $\mu(C) = \mu_n(C) = \lambda^n$. We say that a set $X \subset \mathbf{R}^n$ *has measure zero* if $\forall \varepsilon > 0$, X can be covered by a family of n-cubes the sum of whose measures is less that ε. Thus a countable union of sets of measure 0 is of measure 0. Furthermore if each point of X has a neighbourhood in X of measure 0, then X has measure 0 (Lindelöf).

Lemma 2.6. *Let $U \subset \mathbf{R}^n$ be an open set and $f : U \to \mathbf{R}^n$ a C^1-map. Then if X has measure 0, so does $f(X)$.*

Proof. Each point of X is contained in an open ball $B \subset U \ni$ for some constant $\kappa > 0$

$$\|f(x) - f(y)\| \leq \kappa\|x - y\| \qquad \forall x, y \in B\ .$$

Hence if $C \subset B$ is a cube of edge λ then $f(C)$ is contained in a cube $C' \ni$ such that $\mu(C') < L^n\mu(C)$ where $L = \sqrt{\lambda}\kappa$.

Now $X = \bigcup_{j=1}^{\infty} X_j$ where each X_j is a subset of a ball B as in the preceding paragraph. For each $\varepsilon > 0$, $X_j \subset \bigcup_k C_k$ where each C_k is a cube and $\sum_k \mu(C_k) < \varepsilon$. Hence $f(X_j)$ has measure 0 hence, $f(X)$ has measure 0. This proves the Lemma.

Now let M be an n-dimensional C^∞-manifold (cf. Appendix 3). A subset $X \subset M$ is said to be *of measure* 0 if for each chart (ϕ, U), the set $\phi(U \cap X) \subset \mathbf{R}^n$ has measure 0. It follows from the preceding lemma that this will be true if $\exists$ atlas of charts with this property.

We note that a cube has positive measure, and hence a set of measure 0 in $\mathbf{R}^n$ cannot contain a cube, and therefore must have empty interior. Thus a closed *subset of measure* 0 in $\mathbf{R}^n$, or on a manifold M must be nowhere dense. Suppose $X \subset M$ is of measure 0 and is σ-compact i.e., $X = \bigcup_{n=1}^{\infty} K_n$ where each K_n is compact then each K_n is nowhere dense hence X is nowhere dense. The complement of X is *residual* i.e., contains the union of a countable family of dense open sets, and is dense (by the Baire category theorem).

Lemma 2.7. *Let M, N be manifolds with dim $M <$ dim$N$. If $f : M \to N$ is a C^1-map then $f(M)$ is nowhere dense.*

Proof. This follows from the fact that $F(M)$ has measure 0.

We define a point $x \in M$ to be a *critical point* for a C^1-map $f : M \to N$ if the linear map $T_x f : M_x \to N_{f(x)}$ is not surjective. Denote by $\sum_f$ the set of critical points of f; $f(\sum_f)$ is the set of *critical values* of f, and $N - f(\sum_f)$ is the set of *regular values* of f. We shall now turn to the Morse Sard theorem and we shall only sketch the proof in the case where f is C^∞; this however, is not the best result.

Theorem 2.8. *(Morse-Sard) Suppose M, N are manifolds of dimensions m, n respectively, and $f : M \to N$ a C^r-map. If $r > \max(0, m-n)$ then $f(\sum_f)$ has measure zero in N. The set of regular values of f is residual and therefore dense.*

Proof. (only for the C^∞ case) We shall consider a *local result* viz. in the case where $f : W \to \mathbf{R}^n$ is a C^∞ map, and W is open in $\mathbf{R}^m$. If $m < n$ then $f(W)$ has measure zero. So we shall suppose $m \geq n$.

Write $f(x) = \big(f_1(x), \ldots, f_n(x)\big)$. We note that the critical set $\sum_f$ can be realised to be the sum of three subsets $\sum^1, \sum^2, \sum^3$, defined as follows. $\sum^1$ is the set of points $p \in \sum_f$ such that $\Delta f_i(p) = 0$ for all differential operators Δ of order $\leq \frac{m}{n}$ and all $i = 1, \ldots, m$. $\sum^2$ is the set of points $p \in \sum_f$ such that $\Delta f_i(p) \neq 0$ for some i and some differential operator of order ≥ 2. $\sum^3$ is the set of points $p \in \sum_f$ such that $\frac{\partial f_i}{\partial x_j}(p) \neq 0$ for some i, j. Then $\sum_f = \sum^1 \cup \sum^2 \cup \sum^3$. Each of $f(\sum^1), f(\sum^2)$ and $f(\sum^3)$ has measure 0.

The Morse-Sard theorem is used to show that *transversal* mappings form a dense and open subset of the set $C^r(M, N)$ where M, N are manifolds with C^r fine topology. Let $f : M \to N$ be a C^1-mapping, $A \subset N$ a submanifold. If $K \subset M$, then we write

$f \pitchfork_K A$ to mean that f is *transverse to A along K*, that is to say, if $x \in K$, and $f(x) = y \in A$, then the tangent space N_y is spanned by A_y and the image $(T_x f)(M_x)$. If $K = M$, we write $f \pitchfork A$ (cf. Appendix 3).

Before proceeding to our main theorem on transversal mappings we should observe a few facts which we shall state as Lemmas.

Lemma 2.9. *Let $f : M \to N$ be a C^r map, $r \geq 1$, and $y \in f(M)$ a regular value. Then $f^{-1}(y)$ is a C^r-submanifold of M.*

Proof. It is enough to consider the case where M is an open set in $\mathbf{R}^m$ and N is an open set in $\mathbf{R}^n$, and then the proof follows by using the Inverse Function theorem.

Now suppose $f : M \to N$ is transverse to a submanifold $A \subset N$, i.e., if $f(x) = y \in A$, then $A_y + T_x f(x) = N_y$. Then the following result holds.

Lemma 2.10. *Let $f : M \to N$ be a C^r-map, $r \geq 1$, and $A \subset N$ a C^r-submanifold. If f is transverse to A, then $f^{-1}(A)$ is a submanifold of M, and the co-dimension of $f^{-1}(A)$ in M is the same as the co-dimension of A in N.*

Proof. Let q = codimension of A in N, and $p \ni p + q = n$ = dimension of N. As in the preceding lemma, it is enough to consider $U \times V \subset \mathbf{R}^p \times \mathbf{R}^q$, with $U \times V$ an open neighbourhood of $(0,0) \in \mathbf{R}^p \times \mathbf{R}^q$. Then $f : M \to U \times V$ is transverse to $U \times 0$ if and only if 0 is a regular value for the composite mapping:

$$g : \; M \xrightarrow{f} U \times V \xrightarrow{\pi} V .$$

Now $g^{-1}(0) = f^{-1}(U \times 0)$, and the lemma follows by the preceding lemma.

The next lemma will be stated without proof.

Lemma 2.11.

(a) Let $y_0 \in N$, then the set

$$\left\{ f \in C^r(M,N) \;\middle|\; y_0 \quad \text{is a regular value for} \quad f \right\}$$

is open and dense in $C^r_S(M,N)$, $1 \leq r \leq \infty$.

(b) Suppose $f_0 \in C^r(M,N)$, $y_0 \in N$ is a regular value for f_0, $\mathcal{N}$ is a neighbourhood of f_0 in $C^r_S(M,N)$, and W a neighbourhood of $f^{-1}(y_0)$ in M. Then $\exists g \in N \ni g \pitchfork \{y_0\}$ and $g = f_0$ on $M - W$.

We now define

$$\pitchfork^r_K(M,N,A) \;=\; \left\{ f \in C^r(M,N) \;\middle|\; f \pitchfork_K A \right\} ,$$

and

$$\pitchfork^r(M,N;A) \;=\; \pitchfork^r_M(M,N;A) \,.$$

We shall turn to the following theorem:

Theorem 2.12. *Suppose M, N are manifolds $A \subset N$ a submanifold, and $1 \leq r \leq \infty$. Then*

(a) $\pitchfork^r(M,N;A)$ is residual (hence dense) in $C^r(M,N)$ for the strong topology;

(b) Suppose A is closed in N; if $L \subset M$ is closed then $\pitchfork^r_L(M,N;A)$ is dense and open in $C^r_S(M,N)$.

For the proof we need a "globalisation" lemma. Before stating this lemma we should first explain some preliminary ideas. Let M, N be C^r manifolds, $0 \leq r \leq \infty$. By a C^r *mapping class* we mean a function χ defined as follows. The domain of χ is

the set of triples (L, U, V) with $U \subset M$, $V \subset N$, $L \subset U$ where U, V are open and $L \subset U$ is closed. The mapping χ associates with each such triple, a set of maps $\chi_L(U, V) \subset C^r(U, V)$, and χ further satisfies the following property:

Given triples (L, U, V), a map $f \in C^r(U, V)$ belongs to the class $\chi_L(U, V)$ if $\exists$ triples (L_i, U_i, V_i) and $\exists$ maps $f_i \in \chi_{L_i}(U_i, V_i) \ni L \subset \bigcup_i L_i$ and $f = f_i$ in a neighbourhood of $L_i \quad \forall i$.

We note that the class $\pitchfork^r_L(U, V, V \cap A)$ is a specific example of such a class.

A mapping class is called *rich* if $\exists$ open covers $\mathcal{U}, \mathcal{V}$ of M, N with the property that if $U \subset M$, $V \subset N$ are elements of $\mathcal{U}, \mathcal{V}$ respectively and $L \subset U$ is compact, then $\chi_L(U, V)$ is dense and open in $C^r_W(U, V)$.

The "weak" topology $C^r_W(M, N)$ needs to be defined here. Let $f \in C^r(M, N)$, $(\phi, U), (\psi, V)$ charts on $M, N, K \subset U$ a compact set $\ni f(K) \subset V$, and let $0 < \varepsilon$. Define a neighbourhood $\mathcal{N}^r\big(f; (\phi, U), (\psi, V); \varepsilon\big)$ to be the set of C^r maps $g : M \to N \ni g(K) \subset V$ and $\ni \forall x \in K$ and $\forall$ integers $k \in [0, r]$,

$$\|d^k(\psi f \phi^{-1}(x) - D^k(\psi g \phi^{-1}(x)\| < \varepsilon \, .$$

The topology $C^r_W(M, N)$ is defined to be the one for which sets of the type

$$\mathcal{N}^r\big(f; (\phi, U); (\psi, V); \varepsilon\big)$$

form a subbase.

Lemma 2.13. **(Globalisation)** *Let χ be a rich C^r mapping class on (M, N), $0 \leq r \leq \infty$. Then for every closed set $L \subset M$,*

(a) $\chi_L(M, N)$ is dense and open in $C^r_S(M, N)$,

(b) if L is compact, $\chi_L(M, N)$ is dense and open in $C^r_W(M, N)$.

Lemma 2.14. *Suppose K is a compact set in a manifold U, $\mathbf{R}^p \subset \mathbf{R}^n$ is a linear subspace and V is an open set in $\mathbf{R}^n$. Then*

$$\pitchfork^r_K(U, V; \mathbf{R}^p \cap V)$$

is dense and open in $C^r_W(U, V)$, $1 \leq r \leq \infty$.

For the details we refer the reader to Hirsch [21] pp. 74-77.

This Page Intentionally Left Blank

CHAPTER III

Strong approximation in infinite-dimensional spaces

We shall now turn to some results on strong approximation in infinite dimensional Banach spaces, starting with the theorems of Kurzweil. The concept of C^j-fine approximation in $C^k(U,Y)$ (with $0 \leq j \leq k$) has already been defined (cf. Chapter II §2). Kurzweil's theorems are proved for some of the familiar L^p spaces. Since one of his proofs uses complex-analytic methods, we shall first explain the concept of complex extension of a real Banach space.

§1. Kurzweil's theorems on analytic approximation

Let X be a real Banach space. Then $\tilde{X}$ shall mean the complex space of ordered pairs $z = (x,y) = x + iy$, $x, y \in X$, $i = \sqrt{-1}$ in $\mathbf{C}$, with the usual operations of addition and scalar multiplication and norm

$$\|z\| = \sup_{0 \leq \theta \leq 2\pi} \|(\cos\theta)x - (\sin\theta)y\|$$

Suppose $p(x)$ is a real-valued polynomial defined on X. Then $\exists$ polynomial $\tilde{p}(z)$ defined on $\widetilde{X}$ by the condition: $\tilde{p}(z) = p(z)$ if $z = x + iy$, $y = 0$. The following is the first theorem of Kurzweil.

Theorem 1.1. *([28]) Suppose X is a real reparable Banach space $\ni \exists$ real polynomial $p^*(x)$ on X satisfying*

$$p^*(0) = 0, \qquad \inf_{\substack{x \in X \\ \|x\|=1}} p^*(x) > 0 \,. \tag{1}$$

Let $U \subset X$ be an open set and $f(\cdot) : U \to Y$ be a continuous mapping, Y being a real Banach space. Then $\exists$ analytic mapping $g(\cdot)$ on X, satisfying: $\|g(x) - f(x)\| < 1$ $\forall x \in U$.

Proof. Let $p^*(x)$ be a real polynomial with the properties (1), and suppose $p^*(\cdot)$ is of degree $m > 0$. Then

$$p^*(x) = p_1(x) + p_2(x) + \cdots + p_m(x)$$

where $p_j(\cdot)$ is a homogeneous polynomial of degree j, $1 \leq j \leq m$. Now define

$$p(x) = p_1^2(x) + p_2^2(x) + \cdots + p_m^2(x) .$$

Then $p(x)$ is non negative, and $= 0$ only if $x = 0$. Let $\eta = \inf\{\|p(x)\| \bigm| \|x\| = 1\}$. Then $\eta > 0$.

For $y \in X$, and $r > 0$ define

$$K_r(y) = \{x \in X \bigm| p(x - y) < r\};$$

$$C_r(y) = \{x \in X \bigm| p(x - y) > r\} .$$

Each set $K_r(y)$ is open and bounded. Also for each $r' > 0 \exists r > 0 \ni x \in K_r(0) \Rightarrow \|x\| < r'$. Now for each $x \in U$ $\exists r_x > 0 \ni K_{r_x}(x) \subset U$ and further $\ni$

$$y \in K_{2r_x}(x) \Rightarrow \|f(y) - f(x)\| < \frac{1}{4} .$$

Then $\{K_{r_x}\}_{x \in U}$ is an open covering of U, and because of separability, $\exists$ countable sub-covering $\{K_{r_{x_n}}(x_n)\}_{n=1}^{\infty}$ of U.

Now going back to the proof of Lemma 2.3 in Chapter II we see that $\exists$ open covering $\{D_n\}_{n=1}^{\infty}$ of U which is locally finite and $\ni$ for each $j = 1, 2, \ldots$ $D_j \subset K_{r_{x_j}}(x_j)$.

For convenience we shall again sketch the construction of these open sets D_j. Let $\{\varepsilon_n\}_{n=1}^{\infty}$ be a sequence of positive numbers $\ni$

$$3\varepsilon_i < r_{x_i},\ i = 1,2,\ldots; \qquad 1 > \varepsilon_1 > \varepsilon_2 > \cdots; \qquad \text{and} \qquad \varepsilon_n \to 0 \quad \text{as } n \to \infty .$$

Define

$$\begin{aligned} D_1 &= K_{r_{x_1}}(x_1), \\ D_2 &= C_{r_{x_1}-\varepsilon_2}(x_1) \cap K_{r_{x_2}}(x_2), \\ D_3 &= C_{r_{x_1}-\varepsilon_3}(x_2) \cap C_{r_{x_2}-\varepsilon_3}(x_2) \cap K_{r_{x_3}}(x_3); \ \ldots . \end{aligned}$$

Arguing as in Lemma 2.3 of Chapter II we conclude that these sets form a locally finite open covering of U and $D_j \subset K_{r_{x_j}}(x_j)$ for each $j = 1,2,\ldots$.

We shall need another countable locally finite open covering $\{D_n^*\}_{n=1}^{\infty}$ of $U \ni D_j \subset D_j^* \subset U$ for each $j = 1,2,\ldots$. Define

$$\begin{aligned} D_1^* &= K_{r_{x_1}+2\varepsilon_1}(x_1), \\ D_2^* &= C_{r_{x_1}-3\varepsilon_2}(x_1) \cap K_{r_{x_2}-2\varepsilon_2}(x_2), \\ D_3^* &= C_{r_{x_1}-3\varepsilon_3}(x_1) \cap C_{r_{x_2}-3\varepsilon_3}(x_2) \cap K_{r_{x_3}+2\varepsilon_3}(x_3); \ \ldots . \end{aligned}$$

Then $D_j \subset D_j^* \subset U$ for each $j = 1,2,\ldots$. The set $K_{r_{x_k}-3\varepsilon_\ell}(x_k)$, $\ell > k$ is contained in the complement of the set $D_j^*, j = \ell, \ell+1, \ell+2, \ldots$ and hence $\{D_j^*\}_{j=1}^{\infty}$ is a locally finite open covering of U.

As before we shall denote n-dimensional Euclidean space by $\mathbf{R}^n$. Define the sets $T_j \subset \mathbf{R}^j, j = 1,2,\ldots,$ by:

$$\begin{aligned} T_1 &= \left\{\tau_1 \in \mathbf{R}^1 \;\middle|\; -1 \le \tau_1 \le r_{x_1} + \varepsilon\right\}, \\ T_2 &= \left\{(\tau_1,\tau_2) \in \mathbf{R}^2 \;\middle|\; r_{x_1} - 2\varepsilon_2 \le \tau_1 \le V_2, \quad -1 \le \tau_2 \le r_{x_2} + \varepsilon_2\right\}, \\ T_3 &= \Big\{(\tau_1,\tau_2,\tau_3) \in \mathbf{R}^3 \;\Big|\; r_{x_1} - 2\varepsilon_3 \le \tau_1 \le V_3, \\ &\qquad r_{x_2} - 2\varepsilon_3 \le \tau_2 \le V_3, \quad -1 \le \tau_3 \le r_{x_3} + \varepsilon_3\Big\}; \end{aligned}$$

and so on, where $V_2, V_3, \ldots$ are numbers satisfying:

$$x \in K_{2r_{e_2}}(x_2) \;\Rightarrow\; p(x - x_1) < V_2 - 1 \,,$$

$$x \in K_{2r_{e_3}}(x_3) \;\Rightarrow\; p(x - x_1) < V_3 - 1 \,, \quad p(x - x_2) < V_3 - 1 \,,$$

etc. Now define the numbers $\nu_1, \nu_2, \ldots$ by

$$\frac{1}{\nu_1} = \int_{\mathbf{R}^1} \exp\{-t_1 a_1 \tau_1^2\}\, d\tau_1 = b_1 t_1^{-1/2},$$
$$\frac{1}{\nu_2} = \int_{\mathbf{R}^2} \exp\{-t_2(a_1\tau_1^2 + a_2\tau_2^2)\}\, d\tau_1 d\tau_2 = b_2 t_2^{-1/2},$$

etc., where $b_1, b_2, \ldots$ are positive numbers depending on the constants $a_1, a_2, \ldots$. The positive numbers $a_1, a_2, \ldots,$ and $t_1, t_2, \ldots$ will be chosen later on.

Next we define the functions $\phi_1(z), \phi_2(z), \ldots,$ with $z \in \tilde{X}$, by

$$\phi_1(z) = (\|f(x_1)\| + 1) \cdot \nu_1 \cdot \int_{T_1} \exp\Big\{-t_1 a_1 \big(\tilde{p}(z - x_1) - \tau_1^2\big)\Big\} d\tau_1 \,:$$

$$\begin{aligned}\phi_2(z) = (\|f(x_2)\| + 1) \cdot \nu_2 \cdot \int_{T_2} \exp\Big\{&-t_2 \big[a_1\big(\tilde{p}(z - x_1) - \tau_1\big)^2 \\ &+ a_2\big(\tilde{p}(z - x_2) - \tau_2\big)^2\big]\Big\} d\tau_1 d\tau_2;\end{aligned}$$

$$\begin{aligned}\phi_3(z) = (\|f(x_3)\| + 1) \cdot \nu_3 \cdot \int_{T_3} \exp\Big\{&-t_3 \big[a_1\big(\tilde{p}(z - x_1) - \tau_1\big)^2 \\ &+ a_2\big(\tilde{p}(z - x_2) - \tau_2\big)^2 + a_3\big(\tilde{p}(z - x_3) - \tau_3\big)^2\big]\Big\} \\ &\cdot d\tau_1 d\tau_2 d\tau_3 \;;\end{aligned}$$

and so on. Then the functions $\phi_1(z), \phi_2(z), \ldots$ are analytic in $\tilde{X}$ (because $\phi_j(z)$ is a composition of the mapping

$$z \to \big(\tilde{p}(z - x_1), \tilde{p}(z - x_2), \ldots \tilde{p}(z - x_j)\big)$$

which is analytic in $\widetilde{X}$ and of an analytic function of j complex variables).

Now let $a_n, n = 1, 2, \ldots,$ be positive numbers $\ni$ the series $\sum_{n=1}^{\infty} a_n(1 + \|x - x_n\|)^{4m}$ converges $\forall x \in X$, m being the degree of the polynomial $p^*(z)$; for instance the choice:

$$a_n = \frac{1}{n!}(1 + \|x_n\|)^{4m}$$

will serve the purpose. Then the numbers $t_n, n = 1, 2, \ldots,$ are chosen sufficiently large so that the following three conditions are satisfied ($|T_n|$ being the Lebesgue measure of the set $T_n \subset \mathbf{R}^n$):

$$t_n \geq (n!)^2 \cdot \frac{1}{b_n} \cdot \left(\|f(x_n)\| + 1\right)^2 \cdot |T_n| + 1 \ ;$$

$$|\phi_n(x) - \|f(x_n)\| - 1 \Big| < \frac{1}{2} \ \forall x \in D_n \ ;$$

$$|\phi_n(x)| < \frac{1}{2^{n+3}(\|f(x_n)\| + 1)} \qquad \text{for} \quad x \notin D_n^* .$$

To show that a choice of such numbers t_n is possible, we note that if $x \in D_n$ then for $j = 1, 2, \ldots n-1$: $0 \leq p(x - x_n) < r_{x_n}$; $p(x - x_j) > r_{x_j} - \varepsilon_n$; $p(x - x_j) < V_n - 1$.

If $x \in D_n$ then the open ball in $\mathbf{R}^n$ with centre at the point $\left(p(x - x_1), p(x - x_2), \ldots, p(x - x_n)\right)$ and radius ε_n, is contained in T_n. Also, if $x \notin D_n^*$, then at least one of the inequalities

$$p(x - x_n) < r_{x_n} + 2\varepsilon_n;$$

$$p(x - x_j) > r_{x_j} - 3\varepsilon_n, \qquad j = 1, 2, \ldots, n-1$$

is false. Then by the definition of T_n, it follows that for every $x \notin D_n^*$ the open ball with centre at $\left(p(x - x_1), \ldots p(x - x_n)\right)$ and radius ε_n does not intersect T_n. The choice of the numbers t_n satisfying the above 3 conditions now follows.

Define

$$\phi(x) = \phi_1(x) + \phi_2(x) + \cdots$$

$$H^*(x) = f(x_1)\phi_1(x) + f(x_2)\phi_2(x) + \cdots .$$

The next step is to show that $\phi(\cdot)$ and $H^*(\cdot)$ are analytic. The uniform limit of a sequence of analytic mappings in a complex Banach space is analytic. Hence it is enough to show the following (Hille and Phillips Theorem 3.18.1. [20] p. 113, 1957): for each $x_0 \in U$, $\exists \delta > 0$ and $\exists$ integer $n_0 > 0 \ni \forall z \in \widetilde{X}$ with $\|z\| < \delta$ and $\forall n > n_0$, we have $(\|f(x_n)\| + 1)\ |\phi(x_0 + z)| < \frac{1}{2^n}$.

Let $x_0 \in U$, and let j_0 be the first integer $m \ni x_0 \in K_{r_{*m}}(x_m)$. Then $\exists$ positive number a, and $\exists$ integer $n' > 0 \ni$

$$\tau_{j_0} - p(x_0 - x_{j_0}) > \alpha \qquad \forall \quad \text{points} \quad (\tau_1, \tau_2, \ldots, \tau_n) \in T_n, n > n' \ .$$

We shall find a lower bound for

$$\operatorname{Re} \left\{ \sum_{j=1}^{n} a_j \big(\tilde{p}(x_0 + z - x_j) - \tau_j\big)^2 \right\} .$$

The polynomial $\tilde{p}(x)$ is of degree $2m$; hence

$$p(x_0 - x_j + z) \ = \ p(x_0 - x_j) + Z_j \ ,$$

where $|Z_j| \leq M \cdot \big(1 + \|x - x_j\|\big)^{2m} \cdot \|z\|$, $\|z\| < 1$, M being a positive constant (see Hille and Phillips [20] Theorem 26.2.4 p. 764, 1957). Then

$$\begin{aligned} \big(\tilde{p}(x_0 + z - x_j) - \tau_j\big)^2 \ &= \ \big(p(x_0 - x_j) - \tau_j + Z_j\big)^2 \\ &= \ \big(p(x_0 - x_j) - \tau_j - 2\big(p(x_0 - x_j) - \tau_j\big) Z_j + Z_j^2 \ , \end{aligned}$$

hence

$$\operatorname{Re} \big(\tilde{p}(x_0 + z - x_j) - \tau_j\big)^2 \ \geq \ \Big(\big|p(x_0 - x_j) - \tau_j\big| - |Z_j|\Big)^2 - 2|Z_j|^2 \ .$$

Thus

$$\mathrm{Re}\left\{\sum_{j=1}^{n} a_j\big(p(x_0+z-x_j)-\tau_j\big)^2\right\} \geq$$

$$-2\sum_{j=1}^{n} a_j|Z_j|^2 + a_{j_0}\Big(\big|p(x_0-x_j)-\tau_{j_0}\ \big| - \big|Z_{j_0}\big|^2\Big)$$

$$\geq -2M^2\|z\|^2 \sum_{j=1}^{\infty} a_j\big(1+\|x_0-x_j\|\big)^{4m} + a_{j_0}\big(\alpha - M(1+\|x_0-x_{j_0}\|)^{2m}\|z\|^2\big),$$

where $(\tau_1, \ldots, \tau_n) \in T_n$ and α is positive. The series $\sum_{j=1}^{\infty} a_j(1+\|x_0-x_j\|)^{4m}$ converges; hence $\exists$ positive numbers $\beta, \delta \ni$ the inequality

$$\mathrm{Re}\left\{\sum_{j=1}^{n} a_j\big(p(x_0+z-x_j)-\tau_j\big)^2\right\} > \beta$$

holds $\forall z \in \widetilde{X}$, $\ni \|z\| < \delta$, and $\forall(\tau_1, \ldots, \tau_n) \in T_n, n > n'$. Then

$$\big(\|f(x_n)\|+1\big)\big|\phi_n(x_0+z)\big| \leq \frac{t_n^{n/2}}{b_n}\big(\|f(x_n)\|+1\big)^2 \cdot |T_n| \cdot e^{-\beta t_n}$$

$$\leq \frac{\big(\|f(x_n)\|+1\big)^2}{b_n}|T_n| \cdot \frac{n!}{\beta^n t_n^{n/2}}$$

$$\leq \frac{\big(\|f(x_n)\|+1\big)^2}{b_n}|T_n| \cdot n! \cdot \frac{1}{\beta^n t_n} \leq \frac{1}{n!\beta^n},$$

hence $\exists$ integer $n_0 > 0 \ni$ the inequality

$$\big(\|f(x_n)\|+1\big)|\phi_n(x_0+z)| < \frac{1}{2^n}$$

holds $\forall n > n_0$, and for $\|z\| < \delta$.

Now define

$$H(x) = \frac{H^*(x)}{\phi(x)}.$$

Then $H(\cdot)$ is analytic. Furthermore, let $x \in U$. Then we see that

$$f(x) - H(x) = \frac{1}{\phi(x)} \sum_{i=1}^{\infty} \big\{f(x)\phi_i(x) - f(x_i)\phi_i(x)\big\}.$$

Let I_1 (respectively I_2) denote the set of indices $j \ni x \in D_j^*$ (respectively $x \notin D_j^*$). Then

$$\|f(x) - H(x)\| \leq \frac{1}{\phi(x)} \sum_{j \in I_1} \|f(x) - f(x_j)\| \phi_j(x) + \frac{\|f(x)\|}{\phi(x)} \sum_{j \in I_2} \phi_j(x)$$
$$+ \frac{1}{\phi(x)} \sum_{j \in I_2} \|f(x_j)\| \cdot \phi_j(x) .$$

If $j \in I_1$ then $x \in D_j^* \subset K_{2r_{s_j}}(x_j)$ and $\|f(x) - f(x_j)\| < \frac{1}{4}$. Furthermore $x \in D_\ell$ for some $\ell, \|f(x) - f(x_\ell)\| < \frac{1}{4}$, and $\phi_\ell(x) > \|f(x_\ell)\| + \frac{1}{2}$. Also $\phi(x) \geq \phi_\ell(x) > \|f(x)\|$, $\phi(x) > \frac{1}{2}, \frac{\|f(x)\|}{\phi(x)} < 1$. Then finally

$$\sum_{j \in I_2} \phi_j(x) \leq \sum_{j=1}^{\infty} \frac{1}{2^{j+3}} = \frac{1}{8}, \sum_{j \in I_2} \|f(x_j)\| \phi_j(x) \leq \frac{1}{8} ,$$

and we obtain $\|f(x) - H(x)\| \leq \frac{1}{4} + \frac{1}{8} + \frac{2}{8} < 1$. This completes the proof of Theorem 1.

The next theorem follows form Theorem 1.1.

Theorem 1.2. *([28]) Let X be a Banach space satisfying the hypotheses of Theorem 1.1, $U \subset X$ an open set, $f : X \to Y$ a continuous mapping, Y being a Banach space, and $\varepsilon(\cdot)$ a continuous positive function on U. Then $\exists$ function $g(x)$ which is analytic in U and satisfies: $\|f(x) - g(x)\| < \varepsilon(x) \forall x \in U$.*

Proof. Theorem 1.1 implies that $\exists$ analytic function $\eta(\cdot)$ in U, satisfying: $|\frac{1}{\varepsilon(x)} + 1 - \eta(x)| < 1 \ \forall x \in U$. Then $\eta(x) > \frac{1}{\varepsilon(x)}$. Also by Theorem 1.1, $\exists$ function $H^*(\cdot)$ analytic in U and satisfying:

$$\|\eta(x) f(x) - H^*(x)\| < 1 \qquad \forall x \in U .$$

Let $g(x) = \frac{1}{\eta(x)} H^*(x)$; then $\|f(x) - g(x)\| < \frac{1}{\eta(x)} < \varepsilon(x) \ \forall x \in U$. This proves Theorem 1.2.

Remark. The hypotheses of Theorems 1 and 2 hold in the case of $L^p(\mu)$ or ℓ^p where p is an even positive integer, in this case the p^{th} power of the norm is a polynomial with the required properties, hence Theorems 1.1 and 1.2 are true for these spaces. Likewise the theorems are valid for the Cartesian products of such spaces.

The following corollary of Theorem 1.2 is also a special case of Theorem 1.8 of Whitney (cf. Chapter II, Section I).

Corollary 1.3. *Suppose $U \subset \mathbf{R}^n$ is an open set, $f : U \to \mathbf{R}^1$ a continuous function and $\varepsilon(\cdot)$ a continuous positive function on U. Then $\exists g : U \to \mathbf{R}^1$ which is analytic in G and satisfies: $|f(x) - g(x)| < \varepsilon(x)$ $\forall x \in U$.*

Kurzweils's further results show that in the spaces ℓ^p or $L^p(\mu)$, with $p \geq 1$, $p \neq$ an even integer, as well as in the space $C[0,1]$, uniform approximation to a continuous mapping by an analytic mapping is not possible. We shall now turn to these theorems. However some preliminary definitions and explanations regarding notation are in order.

We shall denote by $\mathbf{R}^+$ the half-line of numbers $\{x \geq 0\}$. Suppose $f(\cdot)$ is a mapping defined on an open set $U \subset X$, X being a real Banach space, with values in a Banach space Y. Let $k > 0$ k a fixed integer. We shall denote by $P(x,h)$ a mapping: $(x,h) \in U \times X \to Y$ such that for fixed $x \in U$, $P(x,h)$ is a polynomial of degree at most k in h and that $P(x,0) = 0$.

Definition. The mapping f is said to be k-times *regularly differentiable* if $\exists$ mapping $P(x,h) : U \times X \to Y$ satisfying:

$$\|f(x+h) - f(x) - P(x,h)\| \leq \alpha(x, \|h\|)\|h\|^k,$$

$\alpha(x,\eta)$ being a nonnegative functional defined on an open subset of $C[0,1] \times \mathbf{R}^+$ and

and satisfying: $\forall x_0 \in U$, $\forall \varepsilon > 0\ \exists \delta > 0 \ni \|x - x_0\| < \delta$ and $0 \le \eta < \delta \Rightarrow \alpha(x,\eta)$ is defined and $0 \le \alpha(x,\eta) < \varepsilon$.

If $k = 1$, then $P(x,h)$ is usually called the *differential of* $f(\cdot)$ and denoted by $\delta f(x,h)$. We note that a regularly differentiable mapping possesses a Fréchet differential and that an analytic mapping is k times regularly differentiable for any $k = 1, 2, \ldots$. Then the following theorem holds.

Theorem 1.4. Kurzweil [28]) *Consider the open ball* $B_R(0) \subset C[0,1]$. *Let* $f : B_R(0) \to Y$ *be (once) regularly differentiable, with values in a weakly complete Banach space. Suppose* ε, r *are two given positive numbers, with* $r + \varepsilon \le R$. *Then* $\exists x \in C[0,1]$ *satisfying:*

$$r \le \|x\| < r + \varepsilon, \qquad \|f(x) - f(0)\| < \varepsilon \ .$$

Note: Before proceeding with the proof of this theorem, we note that the theorem shows that the function $\|x\|$, $x \in C[0,1]$, is *not* the uniform limit of a sequence of regularly differentiable functions in the open unit ball: $B_1(0) \subset C[0,1]$. For suppose $\exists$ regularly differentiable $f : B_1(0) \to \mathbf{R}^1 \ni \big|f(x) - \|x\|\big| < \frac{1}{4}$. Such a conclusion would contradict the conclusion of Theorem 1.4. For Theorem 1.4 implies that $\exists x \in B_1(0) \ni \frac{3}{4} \le \|x\| < 1$, $|f(x) - f(0)| < \frac{1}{4}$; but the inequality $\big|f(x) - \|x\|\big| < \frac{1}{4}$ implies that $|f(0)| < \frac{1}{4}$, and $\big|f(x) - \|x\|\big| < \frac{1}{4}$, and this leads to a contradiction. Hence the function $\|x\|$ on $B_1(0) \subset C[0,1]$ cannot be uniformly approximated by a regularly differentiable function.

Proof of Theorem 1.4. Let $x \in B_R(0) \subset C[0,1]$, and let $V(x)$ be the set of positive numbers η with the property: $\exists$ open subset $H(x;\eta) \subset C[0,1] \times \mathbf{R}^+$ containing all

points (x,ξ) with $0 \le \xi < \eta$, and $\ni$ if $(x',\xi') \in H(x,\eta)$, then $\alpha(x',\xi')$ is defined and $\alpha(x',\xi') < \frac{\varepsilon}{2}$.

Define:

$$\beta(x) = \sup_{\eta \in V(x)} \eta \; ;$$
$$\gamma(x) = \min\bigl(\beta(x), \varepsilon\bigr) \; .$$

Then $\gamma(x)$ has the properties proved in the next lemma.

Lemma 1.5. *$\gamma(\cdot)$ is positive and lower semi- continuous on $B_R(0) \subset C[0,1]$.*

Proof of Lemma. We shall show that $\beta(\cdot)$ is l.s.c. Let $x_0 \in C[0,1]$, and $\xi \in \bigl(0, \beta(x_0)\bigr)$. The set $H(x_0,\xi')$ is open in $C[0,1] \times \mathbf{R}^+$, and contains all points (x_0,ζ) with $0 \le \zeta < \xi'$. The interval $[0,\xi]$ is a compact set hence $\exists$ open set $U \subset C[0,1]$ containing x_0 and $\ni$ all the points (x,ζ), $x \in U, 0 \le \zeta \le \xi$, are contained in $H(x,\xi')$ and $\|x\| < R$ if $x \in U$. Then $\beta(x) \ge \xi \forall x \in U$, and $\beta(\cdot)$ is l.s.c. at x_0. If follows that $\gamma(\cdot)$ is also l.s.c. This proves the lemma.

Before turning to the next lemma, we should state some definitions. Let $q(\cdot)$ be a homogeneous polynomial of degree 1 defined on $C[0,1]$ with values in a weakly complete Banach space (cf. Hille and Phillips [20] Theorem 26.2.4 p. 765). Let $\varepsilon' > 0$, and denote by $T(\varepsilon',q)$ the set of numbers $t \in [0,1]$ with the property: U is an open interval containing the point t, the $\exists$ $x(\cdot) \in C[0,1]$ satisfying: $x(\tau) = 0$ if $\tau \notin U$, $\|x\| < 1, \|q(x)\| > \varepsilon'$.

Lemma 1.6. *The set $T(\varepsilon',q)$ is finite.*

Proof of Lemma. Suppose $T(\varepsilon',q)$ is infinite. Then $\exists$ numbers $t_n \in T(\varepsilon',q)$, $\exists$ open

intervals $U_n \subset [0,1]$, and $\exists$ functions $x_n(\cdot) \in C[0,1]$ for $n = 1, 2, \ldots,$ and satisfying

$$t_n \in U_n; \quad U_i \cap U_j = \phi \quad \text{if} \quad i \neq j; \quad \|q(x_n)\| > \varepsilon'; \quad x_n(\tau) = 0 \quad \text{if} \quad \tau \notin U_n;$$

$$\text{and} \qquad \|x_n\| \leq 1 \; .$$

We now invoke the following theorem of Orlicz:

Lemma 1.7. *(Orlicz [42] Theorem 3, p. 247). Suppose $\{\bar{x}_n\}_{n=1}^{\infty}$ is a sequence of elements in a weakly complete Banach space; and suppose $\exists$ positive number $K \ni$ $\|\bar{x}_{i_1} + \bar{x}_{i_2} + \cdots + \bar{x}_{i_k}\| \leq K$ for every finite sequence $1 \leq i_1 < i_2 < \cdots < i_k$ of integers, $k = 1, 2, \ldots$; then the series $\sum \bar{x}_n$ converges unconditionally.*

To apply this theorem of Orlicz, let $\bar{x}_n = q(x_n)$, and $K = \|q\|$. Then by the preceding lemma (Orlicz) $\sum \bar{x}_n$ converges; this however contradicts the assumption that $\|\bar{x}_j\| = \|q(x_j\| > \varepsilon'$, $j = 1, 2, \ldots$. Thus the Lemma is proved.

We shall now continue with the proof of Theorem 1.4. We shall define a special sequence $\{x_n\} \in C[0,1]$, with $x_1 = 0$, with the following properties. If x_n has been defined with the properties:

$$\|f(x_n)\| \leq \varepsilon \|x_n\|, \qquad \|x_n\| < r \; ,$$

then we shall choose x_{n+1} satisfying:

$$\begin{aligned} \gamma(x_n) &> \|x_{n+1} - x_n\|, \\ \frac{1}{2}\gamma(x_n) &\leq \|x_{n+1}\| - \|x_n\|, \\ \|f(x_{n+1})\| &\leq \varepsilon \|x_n\|, \end{aligned}$$

and thereafter we shall show that such a sequence must be finite. We shall suppose $\varepsilon < 1$.

Let $t' \in [0,1] \ni t' \notin T\ \left(\frac{1}{8}\varepsilon\gamma(x_n), P(x_n,x)\right)$, and $\ni\ |x_n(t')| > \|x_n\| - \frac{1}{4}\gamma(x_n)$. Here $P(x_n,x)$ is the differential of the mapping f. Then $\exists$ open interval U containing $t' \ni$ $\|P(x_n,y)\| \le \frac{1}{8}\varepsilon\gamma(x_n)$ provided y satisfies: $|y(t)| \le 1\ \forall t \in [0,1]$ and $y(t) = 0$ if $t \notin U$. Now consider a fixed function $y(\cdot) \in C[0,1]$, and suppose

$$|y(t)| \le \frac{3}{4}\gamma(x_n) \qquad \text{if} \qquad t \in [0,1],$$

$$y(t) = 0 \qquad \text{if} \qquad t \notin U\ .$$

Let $x_{n+1} = x_n + y$. Then the above 3 required conditions on the sequence $\{x_n\}$ are satisfied because

$$\begin{aligned}\|x_{n+1}\| \ \ge\ |x_{n+1}(t')| \ &=\ |x_n(t')| + \frac{3}{4}\gamma(x_n)\\ &\ge \|x_n\| + \frac{1}{2}\gamma(x_n),\end{aligned}$$

and then by the properties assumed regarding the mapping f,

$$\begin{aligned}&\|f(x_{n+1})\| \le \|f(x_n)\| + \|P(x_n,y)\| + \alpha(x_n,\|y\|)\\ &\le\ \varepsilon\|x_n\| + \frac{1}{8}\varepsilon\gamma(x_n) + \frac{1}{2}\varepsilon\frac{3}{4}\gamma(x_n) = \varepsilon\big(\|x_n\| + \frac{1}{2}\gamma(x_n)\big) \le \varepsilon\|x_{n+1}\|\ .\end{aligned}$$

There must be an element $x_n \ni r \le \|x_n\| < r + \varepsilon$; for otherwise there would be an infinite sequence of points x_n satisfying the above 3 conditions and also $\|x_n\| < r, n = 1,2,\ldots$. Then the series $\sum_{n=1}^{\infty} \gamma(x_n)$ must converge, hence $\{x_n\}$ mush be a Cauchy sequence, $x_n \to x$ say $\|x\| \le r < R$. However the conclusions: $\gamma(x_n) \to 0$ and $\gamma(x) > 0$, contradict the fact that $\gamma(\cdot)$ is lower semicontinuous; this contradiction completes the proof of Theorem 1.4.

We shall now turn to the spaces $\ell^p, L^p(\mu)$, for $p \ge 1$, $p \ne$ an even integer. The methods applied in Theorem 1.4 can be applied again, with slight modifications.

Let $p \ge 1$, $p \ne$ an even integer, be a fixed number, and let $\bar{p}$ be the least integer $\ge$

p. The space X is now ℓ^p, or $L^p(\mu)$. The main result for these spaces is Theorem 1.12 (Kurzweil [28]) below. This result will be proved as a consequence of the next theorem.

Theorem 1.8. *(Kurzweil [30]) Suppose R, r, ε be positive numbers with $R \geq r + \varepsilon$ and suppose the functional f is defined on the open ball $B_R(0) \subset \ell^p$, $p \geq 1, \neq 2, 4, 6, \ldots$, and has the property: $\exists$ functional $w(x,h)$ defined for $x \in B_R(0) \subset \ell^p$, and $h \in \ell^p$, such that for fixed $h, w(x,h)$ is a polynomial of degree at most p and satisfies*

$$|f(x+h) - f(x) - w(x,h)| \leq \alpha(x, \|h\|)\|h\|^p .$$

Here $\alpha(x,\eta)$ is supposed to be defined on an open subset of $B_R(0) \times \mathbf{R}^+$, and is assumed to have the property: $\forall x_0 \in B_R(0)$, and $\forall \varepsilon' > 0$ $\exists \delta > 0 \ni \|x - x_0\| < \delta$ and $0 \leq \eta < \delta \Rightarrow \alpha(x,\eta)$ is defined and satisfies: $0 \leq \alpha(x,\eta) < \varepsilon'$.

Then $\exists x \in \ell^p$ satisfying:

$$r \leq \|x\| < r + \varepsilon, \qquad \text{and} \qquad |f(x) - f(0)| \leq \varepsilon\|x\| .$$

For the proof of Theorem 1.8 the following lemmas are needed.

Lemma 1.9. *Let $v(x)$ be a real homogeneous polynomial in ℓ^p, of degree $v < p$. Let $y = (y_1, y_2, \ldots, y_n, 0, 0, \ldots)$. Then*

$$v(y) = \sum_{1 \leq i_1 < i_2 < \cdots i_v \leq n} a_{i_1 i_2 \ldots i_v} y_{i_1} y_{1_2} \cdots y_{i_v}$$

(here the numbers $a_{i_1 i_2 \ldots i_v}$ do not depend on n and are defined uniquely). Then as $j \to \infty$,

$$\begin{array}{lcl} a_{i_1, i_2, \ldots, i_{v-1}, j} & \longrightarrow & 0, \\ a_{i_1, i_2, \ldots, i_{v-1}, j, j} & \longrightarrow & 0, \\ \vdots & & \\ a_{j,j \ldots, j} & \longrightarrow & 0. \end{array}$$

Proof. This lemma is proved by induction. It holds for $v = 1$. Suppose it holds for polynomials of degree $v-1$, $v < p$. Let

$$\begin{aligned} h_1 &= \{1, 0, 0, \ldots\}, \\ h_j &= \left\{ \underbrace{0, \ldots, 0}_{j-1}, 1, 0, \ldots \right\}, \qquad j = 2, 3, \ldots . \end{aligned}$$

Then the differentials

$$\delta V(y, h_i) = \frac{\partial}{\partial y_i} \sum_{i \le i_1 < i_2 < \cdots < i_v \le n} a_{i_1, i_2, \ldots, i_v} y_{i_1} y_{i_2} \cdots y_{i_v} \quad (1 \le i \le n)$$

are homogeneous polynomials of degree $v-1$ in y (cf. Hille and Phillips [20] Theorem 26.2.9, p. 765). Hence the assertions of the lemma (except the last one) clearly hold; and if the last assertion were not true, then $\exists$ sequence $j_1, j_2, \ldots$ with the property: let

$$\begin{aligned} e_k &= \{a_{k,1}, a_{k,2}, \ldots\} \quad k = 1, 2, \ldots, \quad \text{where} \\ a_{k,j_1} &= a_{k,j_2} = \ldots = a_{k,j_k} = \frac{1}{k^{1/p}}, \\ a_{k,j} &= 0 \quad \text{if} \quad j \neq j_1, j_2, \ldots, j_k; \end{aligned}$$

then $V(e_k) \to \infty$ as $k \to \infty$. This contradiction establishes the lemma.

This lemma in turn implies the next one.

Lemma 1.10. *Let $\underline{p}$ be the greatest integer less than or equal to p, and let ε, Q be two positive numbers. Suppose $q(\cdot)$ is a real-valued polynomial on ℓ^p, of degree at most $\underline{p}$. Then $\exists x \in \ell^p$ satisfying*

(1) $$\|x\| = Q\,,$$

(2) $$|q(x) - q(0)| < \varepsilon$$

(3) x *has only a finite number of coordinates different from 0.*

Proof of Lemma. If p is not an integer, the proof is immediate from the last assertion (for v) of the earlier lemma. Suppose p is an integer, then p must be odd. Then let $q(x) = q_1(x) + q_2(x)$ where $q_1(\cdot)$ is of degree at most $p-1, q_2(\cdot)$ is homogeneous of degree p. So $q_2(x)$ is a continuous odd functional. Thus the lemma is proved.

Proof of Theorem 1.8. *This consists in an almost verbatim repetition of the proof of Theorem 1.4.*

As is easily verified, every polynomial satisfies the conditions of Theorem 1.8. Hence we obtain the next corollary.

Corollary 1.11. *Let ε, r_1, r_2 be the positive numbers, $r_1 < r_2$. Suppose $q(\cdot)$ is a real-valued polynomial in $\ell^p(p \geq 1, \neq 2, 4, 6, \ldots)$. Then $\exists x \in \ell^p$ satisfying*

(1) $$r_1 < \|x\| < r_2 \ ,$$

(2) $$|q(x) - q(0)| < \varepsilon,$$

(3) *the point x has only a finite number of coordinates different from 0.*

Proof of Corollary. Theorem 1.8 implies that $\exists y \in \ell^p$ satisfying the assertions (1) and (2) The polynomial $q(\cdot)$ is a continuous functional. Hence $\exists$ point $x \in \ell^p$ satisfying all the assertions (1)-(3).

We shall now turn to the crucial theorem.

Theorem 1.12. *([30]). Let ε, R, r be positive numbers with $R \geq r + \varepsilon$. Suppose the functional $f(\cdot)$. is defined in the open ball $B_R(0) \subset X = \ell^p$ $(p \geq 1, p \neq 2, 4, 6, \ldots)$, and is $\bar{p}$ times regularly differentiable. Then $\exists$ $x \in X$ satisfying*

$$r \leq \|x\| < r + \varepsilon, \quad |f(x) - f(0)| \leq \varepsilon \|x\| \ .$$

Proof of Theorem 1.12. The proof of Theorem 1.12 for $X = \ell^p$ consists again in an almost verbatim repetition of the proof of Theorem 1.8, except that Corollary 1.11 is now used instead of Lemma 1.10. Thus the theorem follows in the case $X = \ell^p$.

Furthermore, $L^p(\mu)$ (for the given values of p), contains a subspace isometric to ℓ^p. Hence Theorem 1.12 follows for $X = L^p(\mu)$.

Remark. For a different proof of Theorem 1.12, see [62]. Going back to Kurzweil's Theorem 1.1, the condition

(A) $\exists$ real-valued polynomial $q^*(x)$ on X satisfying

$$q^*(0) \;=\; 0, \quad \inf_{\|x\|=1} q^*(x) > 0$$

can be shown to be also necessary for uniform analytic approximation to a continuous mapping on X, if X is uniformly convex. In fact the following theorem has been proved.

Theorem 1.13. *([31]) Suppose X is a uniformly convex Banach space. Then the following three conditions are equivalent:*

(A) $\exists$ real-valued polynomial $q^*(x)$ in X satisfying:

$$\inf_{\|x\|=1} \left|q^*(x) - q^*(0)\right| > 0 \;;$$

(C): if $U \subset X$ is an open set, $f(\cdot) : U \to Y$ is a continuous mapping where Y is a Banach space, and if ε is a positive number, then $\exists$ mapping $g : U \to Y$ such that $g(\cdot)$ is analytic and satisfies:

$$\|g(x) - f(x)\| < \varepsilon \qquad \forall x \in U \,.$$

For reasons of brevity, however, we shall refer the reader to Kurzweil's original paper ([31]), and end here our account of his work.

§2. Smoothness properties of norms in L^p-spaces

We shall briefly recall the smoothness properties of the norms in the classical L^p-spaces. Questions of strong approximation by smooth functions in these spaces are related to the smoothness properties of the norms. For basic information about this topic we refer the reader to the monograph [27] Vol. I of Köthe (Chapter V §26, pp. 342ff.). Further properties are derived at length in Bonic and Frampton [5], and also [58]. Some notation should first be explained.

Definition. A Banach space X is UF^k*-smooth* if the norm in X is k-times uniformly Fréchet differentiable over the unit sphere $S_X = \{x \in X \big| \|x\| = 1\}$. X is BF^k*-smooth* if the norm in X is k-times Frechet differentiable at $x \neq 0$ and if further the map $x \to T_x^k, T_x^k$ being the kth successive Fréchet derivative of the norm $\|x\|$ on $X - \{0\}$, is bounded on the unit sphere S_X, i.e., $\sup\{\|T_x^i\| x \in S_X\} < \infty$ for $1 \leq i \leq k$.

Notation. We shall denote by $R(\lambda, \mu)$ an annular region of the type $\{x \in X \big| \lambda < \|x\| < \mu\}$ for $0 < \lambda < \mu$.

Theorem 2.1. *(cf. [58]) A Banach space X is UF^k-smooth $\Leftrightarrow$ the norm $\| \quad \|$ in X is uniformly k-times continuously differentiable in arbitrary annular regions $R(\lambda, \mu)$. A Banach space which is UF^k-smooth is also UF^i-smooth for $1 \leq i \leq k$, as also BF^k-smooth.*

Let (Ω, Σ, μ) be an arbitrary measure space with μ a positive measure. μ is assumed to be non-trivial, and we shall consider $1 \leq p < \infty$; so $L^p(\mu)$ is of dimension ≥ 2. When Ω is the set of positive integers and each point is an atom of mass 1, we write ℓ^p instead of L^p. For a positive number p, denote by $E(p)$ the integral part of p.

The main properties of $L^p(\mu)$-norms are proved in the next theorem.

Theorem 2.1. *(See [5], [58]) The norm in $L^p(\mu)$ is*

(1) uniformly $(p-1)$ times differentiable over every region $R(\lambda, \mu)$, if p is an odd integer;

(2) uniformly k-times continuously differentiable for every positive integer k, over every region $R(\lambda, \mu)$, if p is an even integer; and

(3) uniformly $E(p)$-times continuously differentiable over every region $R(\lambda, \mu)$ if p is not an integer.

We shall follow [5] for the proof.

Proof. We shall write $\alpha(x) = \|x\|^p = \int |x(\omega)|^p \mu(d\omega)$. Let n denote the greatest integer strictly less than p, and define $\phi \in C^n(\mathbf{R}^1, \mathbf{R}^1)$ by: $\phi(t) = |t|^p$. Then $D^k\phi(t) = p(p-1)\dots(p-k+1)|t|^{p-k}(\text{sgn } t)^k$, where $\text{sgn}(t) = +1$ if $t > 0, -1$ if $t < 0$, and $= 0$ if $t = 0$. An application of Taylor's formula gives us:

$$\phi(t+h) - \sum_{k=0}^{n} D^k\phi(t)\frac{h^{(k)}}{k!} = \int_0^1 (1-s)^{n-1}\Big\{D^n\phi(t+sh) - D^n\phi(t)\Big\}\frac{h^{(n)}}{(n-1)!}ds .$$

The specific form of the function $D^n\phi$ on $t \geq 0$ shows that

$$\left|\phi(t+h) - \sum_{k=0}^{n} D^k\phi(t)\frac{h^{(k)}}{k!}\right| \leq p(p-1)\dots(p-n+1)\frac{|h|^p}{n!} .$$

For any $x(\cdot) \in L^p$, and $\forall k = 0, 1, \dots, n$, let $A_k(x)$ denote the continuous k-linear functional on L^p, defined by $A_k(x)(h_1, \dots, h_k) = \int_X D^k\phi(x(\omega)).\ h_1(\omega)h_2(\omega)\dots h_k(\omega)\mu(d\omega)$, $h_1, \dots, h_k \in L^p$. The continuity of $A_k(x)$ follows by an application of

Hölder's inequality. Then

$$
\begin{aligned}
&\left|\alpha(x+h) - \sum_{k=0}^{n} \frac{1}{k!} A_k(x)(h, \ldots, h)\right| \\
&\leq \int_X \left|\phi\big(x(\omega)+h(\omega)\big) - \sum_{k=0}^{n} D^k\phi\big(x(\omega)\big)\big(h(\omega)\big)^{(k)}\right| \mu(d\omega) \\
&\leq \frac{p(p-1)\ldots(p-n+1)}{n!} \cdot \int_X \left|h(\omega)\right|^n \left|h(\omega)\right|^{p-n} \mu(d\omega) \\
&\leq \frac{p(p-1)\ldots(p-n+1)}{n!} \|h\|^n \|h\|^{p-n} .
\end{aligned}
$$

Thus (cf. Lusternik and Sobolev [36] §42) we obtain that $D^k\alpha(x) = A_k(x), k = 0, 1, \ldots, n$.

Suppose p is an even integer. Then

$$\big(\mathrm{sgn}\ x(\omega)\big)^n \cdot |x(\omega)|^{p-n} = x(\omega) ,$$

and

$$D^n\alpha(x)(h_1, \ldots, h_n) = p(p-1)\ldots(p-n+1) \int_X x(\omega)h_1(\omega)\ldots h_n(\omega)\mu(d\omega)$$

which shows that $D^n\alpha$ is linear in x. Hence $D^{n+1}\alpha$ is constant and $D^{n+j}\alpha = 0$ for $j \geq 2$, i.e., α is C^∞.

Suppose p is not an even integer. Then

$$
\begin{aligned}
&\left|\big\{D^n\alpha(x+h) - D^n\alpha(x)\big\}(h_1, \ldots, h_n)\right| \\
&= p(p-1)\ldots(p-n+1) \left| \int_X \Big(|x(\omega)+h(\omega)|^{p-n} \Big\{\mathrm{sgn}\ \big(x(\omega)+h(\omega)\big)\Big\}^n \right. \\
&\qquad \left. - \left|x(\omega)\right|^{p-n} \big\{\mathrm{sgn}\ \big(x(\omega)\big)\big\}^n \Big) h_1(\omega)\ldots h_n(\omega)\mu(d\omega) \right| \\
&\leq p(p-1)\ldots(p-n+1) \left| \int_X |h(\omega)|^{p-n} |h_1(\omega)\ldots h_n(\omega)| \mu(d\omega) \right. \\
&\leq p(p-1)\ldots(p-n+1) \|h\|^{p-n} \|h_1\| \ldots \|h_n\| ,
\end{aligned}
$$

and therefore

$$\|D^n\alpha(x+h) - D^n\alpha(x)\| \leq \ p(p-1)\dots(p-n+1)\|h\|^{p-n}$$

which shows that α is C^n.

The uniformity part of the assertion of the theorem is also clear from the arguments, and the proof is complete.

§3. C^∞ partitions of unity in Hilbert space

The first theorem of Kurzweil (cf. Theorem 1.1 in this chapter) depends on complex analytic methods; arguments based on partitions of unity are not used there. Partitions of unity provide a very elegant tool, and we shall present in the following sections some interesting results derived with the help of this tool.

J. Eells has shown the existence of C^∞-partitions of unity on a separable Hilbert manifold (for a proof see Lang [32]). Similar C^∞ partitions of unity were constructed by Torunczyk and K. Sundaresan (cf. [58] for an account).

We shall sketch below the proof of J. Eells showing the existence of C^∞ partitions of unity in a separable Hilbert space; for complete details the reader should refer to [32] p. 37.

Proposition 3.1. *A separable Hilbert space admits C^∞ partitions of unity.*

Sketch of proof. Let U be any nonempty open set in a separable Hilbert space $\mathcal{H}$. Suppose $\mathcal{W} = \{B_i\}_{i=1}^\infty$ is a countable covering of U by open balls, where $B_i = B_{r_i}(x_i), i = 1, 2, \ldots$. Then proceeding as in Chapter II, §2, we let $\{\varepsilon_i\}_{i=1}^\infty$ be a sequence of positive numbers $\ni$

$$3\varepsilon_i < r_i, 1 > \varepsilon_1 > \varepsilon_2 > \cdots > 0, \qquad \text{and} \qquad \varepsilon_i \to 0 \qquad \text{as} \quad i \to \infty .$$

Then define: $V_1 = B_1, V_2 = B_2 \cap \mathcal{C}B_{r_1-\varepsilon_2}$, and generally,

$$V_j = B_j \cap \mathcal{C}\overline{B_{r_1-\varepsilon_b}(x_1)} \cap \mathcal{C}\overline{B_{r_2-\varepsilon_j}(x_2)} \cap \ldots \cap \mathcal{C}\overline{B_{r_{j-1}-\varepsilon_j}(x_{j-1})} \,;$$

then $U = \bigcup\limits_{j=1}^{\infty} V_j, V_j \subset B_j$ for $j = 1, 2, \ldots,$ and the covering $\mathcal{V} = \{V_j\}_{j=1}^\infty$ is locally finite.

We now need the following lemma.

Lemma 3.2. *(cf. [32], p. 36) Suppose $B, B_1, \ldots, B_m$ are open balls in H; let V be the "scallopped" set:*

$$V = B \cap \mathcal{C}\bar{B}_1 \cap \ldots \cap \mathcal{C}\bar{B}_m .$$

Then exists C^∞ function $\theta(\cdot) : \mathcal{H} \to [0,1] \ni \theta(x) > 0$ and $\theta(\cdot) \equiv 0$ on $\mathcal{C}V$.

Proof. Let $\phi, \phi_1, \ldots, \phi_m$ be C^∞ functions form $\mathcal{H} \to [0,1] \ni$

$$\phi(x) > 0 \quad \text{if} \quad x \in B, \quad \text{and} \quad \phi(\cdot) \equiv 0 \quad \text{on} \quad \mathcal{C}B ,$$

and

$$\text{for} \quad i = 1, \ldots, m, \phi_i(x) \equiv 1 \quad \text{if} \quad x \in \bar{B}_i \qquad \text{and} \quad 0 \le \phi_i(x) < 1 \quad \text{if} \quad x \in \mathcal{C}\bar{B}_i .$$

Then define $\theta(x) = \phi(x)\big(1 - \phi_1(x)\big)\ldots\big(1 - \phi_m(x)\big)$. This function $\theta(\cdot)$ is the C^∞ function which we are looking for.

Proof of Theorem 3.1. For each $i = 1, 2, \ldots,$ let θ_i be the C^∞ function from $\mathcal{H}$ to $[0,1] \ni \theta_i(x) > 0$ if $x \in V_i$ and $\theta_i(\cdot) \equiv 0$ on $\mathcal{C}V_i$. Now define $\lambda_i = \frac{\theta_i}{\Sigma_i \theta_i}$, $i = 1, 2, \ldots$. Then $\{\lambda_i\}_{i=1}^{\infty}$ is the C^∞-partition of unity subordinate to $\mathcal{V}$.

§4. Theorem of Bonic and Frampton

The theorems of this section and the next one serve as good illustrations of the use of partitions of unity. These were originally formulated for Hilbert manifolds. We shall however give the proof only for an open set in a separable Hilbert space.

Theorem 4.1. *(cf. [5]) Suppose U is a nonempty open set in a separable Hilbert space $\mathcal{H}$, and $f \in C^0(U,F)$ where F is a Banach space. Let $\varepsilon(\cdot)$ be a continuous positive function on U. Then $\exists g(\cdot) \in C^\infty(U,F) \ni \|g(x)-f(x)\| < \varepsilon(x)\ \forall x \in U$. In other words $C^\infty(U,F)$ is dense in $C^0(U,F)$ in the C^0-fine topology on $C^0(U,F)$.*

Proof. Let V be an open subset of U. Let $\varepsilon > 0$ and let $\mathcal{V} = \{V_\alpha\}_{\alpha\in I}$ be an open covering of F by balls of radius $\varepsilon/2$. Then $\{U_\alpha\}_{\alpha\in I}$ where $U_\alpha = f^{-}1(V_\alpha)$ is an open covering of V. We express each U_α as a union of open balls $\{B_{\alpha,i}\}_{i\in I_\alpha}$, and therefrom obtain a countable covering $\mathcal{D} = \{D_n\}_{n=1}^\infty$ of V by open balls $D_n \ni f$ maps each D_n into a set of diameter $\le \varepsilon$. Now let $\{\psi_n\}_{n=1}^\infty$ be a C^∞-partition of unity subordinate to $\mathcal{D}$, where w.l.g. we suppose that no $\psi_n \equiv 0$. For each n let $x_n \in V \ni \psi_n(x_n) > 0$. Then $\psi_n(x) > 0 \Rightarrow \|f(x_n) - f(x)\| < \varepsilon$. Then $\Sigma f(x_n)\psi_n(x) \in C^\infty$ and $\|\Sigma f(x_n)\psi_n(x) - f(x)\| < \varepsilon \forall x \in V$.

Next, for each $x \in U \exists$ open ball $B_{r_x}(x) \subset U \ni y \in B_{r_x}(x) \Rightarrow \frac{\varepsilon(x)}{2} < \varepsilon(y)$. Separability $\Rightarrow \exists$ countable covering $\{B_{r_{x_n}}(x_n)\}_{n=1}^\infty$ of U where for each $n = 1,2,\ldots$, $y \in B_{r_{x_n}}(x_n) \Rightarrow \frac{\varepsilon(x_n)}{2} < \varepsilon(y)$. For convenience we shall write r_n for $r(x_n)$. By the arguments of the previous paragraph $\exists g_{x_n} \in C^\infty\big(B_{r_n}(x_n),F\big) \ni \forall y \in B_{r_n}(x_n)$, $\|g_{x_n}(y) - f(y)\| < \frac{\varepsilon(x_n)}{2} < \varepsilon(y)$. Now let $\{\lambda_n\}_{n=1}^\infty$ be a C^∞ partition of unity subordinate to the covering $\{B_{r_n}(x_n)\}$. Define $g(\cdot) = \sum \lambda_n(\cdot)g_{x_n}(\cdot)$. Then $g \in C^\infty$ and

satisfies: $\|g(y) - f(y)\| < \varepsilon(y) \ \ \forall y \in U$. This completes the proof.

§5. Smale's Theorem

To obtain an infinite dimensional version of Sard's theorem, the concept of "set of zero measure" has to be first replaced by a more satisfactory one which makes sense even in infinite dimensional spaces. One such concept is "set of first category", and Smale was able to show that Sard's theorem holds if we substitute the concept of first category for zero measure. In this section we shall present this theorem and a consequence of it.

This infinite dimensional Sard's theorem needs a nonlinear concept of Fredholm operator. First we need some definitions and facts concerning linear operators. A *Fredholm operator* is a continuous linear operator $L : X \to Y$ from a Banach space to another, with the properties:

(a) dim ker $L < \infty$;

(b) range (L) is closed;

(c) coker $(L) = Y/\text{range}\,(L)$ has finite dimension.

The *index* of a Fredholm operator $L : X \to Y$ is dim ker $(L)-$dim coker (L); the index is an integer. The following main result should be noted, the proof of which with more details can be found in [15] .

Lemma 5.1. *The set $F(X,Y)$ of Fredholm operators is open in the space $L(X,Y)$ of all bounded linear operators, in the norm topology. The index is a continuous function on $F(X,Y)$.*

The nonlinear generalisations can be conveniently explained using concepts of differentiable Banach manifolds. We shall always assume that these manifolds are con-

nected and have a countable base. Let M, V be two such manifolds. A *Fredholm map* is a C^1-mapping $f : M \to N$ such that for each $x \in M$, the derivative $Df(x)$: $T_x(M) \to T_{f(x)}(N)$ is a (linear) Fredholm operator. The index of f is defined to be the index of $Df(x)$ for some x. Our assumption is that f is C^1 and M is connected; hence the index does not depend on x.

Let $f : M \to N$ be a C^1-map. A point $x \in M$ is called a *regular point* of f if $Df(x) : T_x(M) \to T_{f(x)}(N)$ is surjective, and is *singular* if not regular. The images of singular points under f are called *critical values*, and their complement the *regular values*. Hence if $y \in N$ is not in $f(M)$ then y is automatically a regular value.

The finite dimensional Sard's theorem explained earlier in Chapter II will be needed in the sequel. Henceforth in this section "almost all" shall mean "except for a set of the first category". Smale's main theorem is the following (cf. [58]).

Theorem 5.2. *Let $f : M \to N$ be a C^q-Fredholm map with $q >$ max (index $f, 0$). Then the regular values of f are almost all of N.*

Proof. We recall that M has a countable base. Also a countable union of sets of the first category is again a set of the first category. Hence it is enough to prove the theorem locally. Hence it is enough to assume that the given f is a map: $U \to Y$ where U is an open set in a Banach space X, and Y is another Banach space.

Let $x_0 \in U$, and $A = Df(x_0) : X \to Y$. Now dim Ker $A < \infty$, X can be expressed as $X = X_1 \times$ Ker A where X, is a Banach space, and $x_0 = (p_0, q_0)$, $p_0 \in X_1$, $q_0 \in$ ker A. Then, $\forall (p, q)$ sufficiently close to (p_0, q_0), the first partial derivative $D_1 f(p, q)$: $X_1 \to X$ maps X_1 injectively onto a closed subspace of X. Now use the implicit function

theorem: we can choose a product neighbourhood $B_1 \times B_2$ of (p_0, q_0) in $X_1 \times \text{Ker } A$ such that B_2 is compact, and if $q \in B_2$ then f restricted to $B_1 \times q$ is a differentiable homeomorphism onto its image.

We now need the following lemma. A map ϕ is called *proper* if the inverse image of a compact set, under ϕ, is compact.

Lemma 5.3. *A Fredholm map is locally proper, i.e., if $f : M \to N$ is Fredholm, and $x \in M$, then $\exists$ neighbourhood W of $x \ni f\big|_W$ is proper.*

Proof of Lemma. Choose $N(x) = B_1 \times B_2$ as above and let $f(x_0) = y_i$ tend to y, where $x_i = (p_i, q_i) \in B_1 \times B_2$. It is enough to show that the x_i have a convergent subsequence. Now B_2 is compact, hence we can assume that $q_i \to q$, and since $f(p_i, q) \to y$, we can even assume $q_i = q$. Now f restricted to $B_1 \times q$ is a homeomorphism onto its image. Hence $p_i \to p$. This proves the lemma.

Proof of Theorem 5.2 (cont'd). Let $x_0 \in M$, and again let $B_1 \times B_2 \subset X_1 \times \text{Ker }(A)$ as above. The critical points of f form a closed set. Hence by the preceding lemma it is enough, given a neighbourhood U_1 of $f(x_0)$ in Y to find a regular value of f in U_1.

Let π be the projection: $Y \to Y/\text{Range }(A)$. From the hypotheses of the theorem, Sard's theorem can be applied to the map $\phi : \{p_0\} \times \text{Ker }(A) \to Y/\text{Range }(A)$ defined by: $\phi(q) = \phi \circ f(p_0, q)$ to give a regular value z of ϕ in πU_1. Let $y \in \pi^{-1}(z) \cap U_1$. Then such a y is a desired regular value. This completes the proof of the theorem.

The next theorem (cf. [58] Theorem 3.1) follows as an interesting consequence of the preceding theorem. First some definitions are in order.

Let $f : M \to N$ be a C^1 map, and g: $W \to N$ be a C^1 imbedding. We say that

f is *transversal to* g if for each $(x,y) \in M \times W$ such that $f(x) = g(y)$ the two spaces range $(Df(x))$, range $(Dg(y))$, span the tangent space $T(f(x)(N)$.

Theorem 5.4. *Let $f : M \to N$ be a C^q-Fredholm mapping, and $g : W \to N$ a C^1-imbedding of a finite dimensional manifold W, with*

$$q > \max(\text{index } f + \dim W, 0) .$$

Then $\exists C^1$-approximation g' of g such that f is transversal to g'. Furthermore if f is transversal to $g|_A$ where A is a closed subset of W, then g' may be chosen so that $g' = g$ on A.

Proof. M as well as W, has a countable base. Hence a standard argument reduces the proof of this theorem to the following lemma.

Lemma 5.5. *Let M, W as in the preceding theorem. Then $\exists$ neighbourhood U_1 of y* and for any $\varepsilon > 0$ an εC^1 approximation g' of g such that f is transversal to $g'|_{U_1}$.

Proof of Lemma. We can assume that $\exists$ neighbourhood of $g(y)$ in N, as follows: $U_2 \subset \mathbf{R}^p$, N=Banach space $F = \mathbf{R}^p \times F_1$, and $g : U_2 \to \mathbf{R}^p \times 0$ is identity $\times 0$, $\pi : F \to F_1$ the projection. Let U_1 be a neighbourhood of $y \ni \bar{U}_1 \subset \text{int}(U_2)$, and ϕ a C^∞ function which equals 1 on $U_1, 0$ on ext (U_2). Then by the main theorem, let $z \in F_1$ be close to $0 \ni \pi \circ f$ has z as a regular value on $f^{-1}(g(U_2))$. Now define g' as the translate of g by z on U_1, smoothed by ϕ.

This proves the lemma and hence Theorem 5.4 of Smale.

§6. Theorem of Eels and McAlpin

In this section we shall give an account of the following theorem of J. Eells and J. McAlpin (cf. [10]). However we shall formulate the result for Banach spaces rather than for manifolds.

Suppose X, Y are smooth connected manifolds modelled on Banach spaces, and $\phi : X \to Y$ is a C^1-map. A point $x \in X$ is a critical point of ϕ if the differential $\phi_*(x) : T_xX \to T_{\phi(x)}Y$ is not surjective. The set of critical points of ϕ is denoted by C_ϕ. The mapping ϕ is called *residual* (or of *Sard type*) if $\phi(C_\phi)$ has no interior point in Y.

Remark. The theorem of Smale (Theorem 5.2 in this Chapter) shows that to establish an infinite dimensional analogue of the Morse-Sard theorem, strong restrictions are necessary.

Theorem 6.1. *Suppose that $\mathcal{H}$ is a separable Hilbert space, and F is a Banach space. Then the C^∞-smooth residual maps are dense in the C^0-fine topology on $C^0(\mathcal{H}, F)$.*

Some lemmas are needed for the proof of this theorem.

Lemma 6.2. *Let U be an open subset in $\mathcal{H}$. Then for any closed set $C \subset U$ and neighbourhood V of C in U, $\exists$ countable collection $\{U_i\}_{i=1}^{\infty}$ of open balls in $\mathcal{H}$ such that*

(1) those with even (respectively odd) subscripts are in and cover V (resp., $U - C$), and

(2) the centres a_i of the U_i are linearly independent points in $\mathcal{H}$.

Proof of Lemma 6.2. For each $x \in V$ let $B_{r_x}(x)$ be an open ball contained in V. Then $\exists$ countable collection $\left\{B_{\frac{r_{x_n}}{2}}(x_n)\right\}_{n=1}^{\infty}$ covering V. Similarly let $\left\{B_{\frac{r_{y_n}}{2}}(y_n)\right\}_{n=1}^{\infty}$ be a countable covering of $U - C$ by open balls $B_{\frac{r_{y_n}}{2}}(y_n)$ where each $y_n \in U - C$.

Now proceed by induction: for each $B_{\frac{r_{x_n}}{2}}(x_n)$ (resp. $B_{\frac{r_{y_m}}{2}}(y_m)$) select open balls U_{2n} labelled with even subscripts (resp. U_{2m+1} labelled with odd subscripts) such that $B_{\frac{r_{x_n}}{2}}(x_n) \subset U_{2n} \subset B_{r_{x_n}}(x_n)$ (resp. $B_{\frac{r_{y_m}}{2}}(y_m) \subset U_{2m+1} \subset B_{r_{y_m}}(y_m)$) and whose centres a_i are linearly independent in $\mathcal{H}$. This completes the proof of Lemma 6.2.

Before turning to the next lemma, some notation should be explained. For any subset $A \subset U$, and any $r > 0$, let $(A, r) = \{x \in U \mid d(x, A) < r\}$. Set $A^c = U - A$.

Lemma 6.3. *Let $V_1 = U_1$, and for $k \geq 1$, let $V_{2k+1} = U_{2k+1} \cap (U_1^c, \frac{1}{k}) \cap \ldots \cap (U_{2k-1}^c, \frac{1}{k})$. Then $\{V_{2i+1}\}_{i=0}^{\infty}$ is a locally finite open refinement of the covering $\{U_{2i+1}\}_{i=0}^{\infty}$.*

Also $V_2 \overset{\text{def}}{=} U_2$, and for $k \geq 2$, $V_{2k} \overset{\text{def}}{=} U_{2k} \cap (U_2^c, \frac{1}{k}) \cap \ldots \cap (U_{2k-2}^c, \frac{1}{k})$ determines a locally finite refinement of $\{U_{2i}\}_{i=0}^{\infty}$.

For the proof see Lang [32] p. 32

Lemma 6.4. *Let U be an open set in a separable Hilbert space $\mathcal{H}$. Then for each pair of disjoint closed subsets $C_0, C_1 \subset U$ $\exists$ μ_1-residual function $\phi : U \to [0,1]$ with $\phi^{-1}(0) = C_0, \phi^{-1}(1) = C_1$.*

Proof of Lemma 6.4. Take $C = C_1$ and $V = U - C_0$ in Lemma 6.2. Then the composition of a C^∞-function $r_{ij} : \mathbf{R}^1 \to \mathbf{R}^1$ and $\|x - a_j\|$ gives C^∞-smooth functions $f_{ij} : U \to \mathbf{R}^1$ for $j \leq i$ and $(-1)^j = (-1)^i$ such that

(1) f_{ii} is strictly positive on U_i and $= 0$ elsewhere;

(2) for $j < i$ we have $f_{ij} = 1$ outside $U_j = B_{r_j}(a_j)$, $f_{ij} = 0$ on $B_{r_j - \frac{1}{i}}(a_j)$; and $0 < f_{ij}(x) < 1$ in between.

Now use the Hilbert space structure on $\mathcal{H}$; we find that the gradient $\nabla f_{ij}(x) = \alpha_{ij}(\|x - a_j\|)(x - a_j)$ for C^∞-real functions α_{ij}; we can assume $\alpha_{ii}(t) = 0$ only

when $t = 0$ or $t \geq r_i$. Let $f_i(x) = \Pi f_{ij}(x)$; then $f_i > 0$ on $V_i, = 0$ elsewhere, and $\nabla f_i(x) = \sum \beta_{ij}(x)(x - a_j)$ for suitable real functions β_{ij}. Furthermore $\beta_{ii}(x) = \alpha_{ii}(\|x - a_i\|) \prod_{j<i} f_{ij}(x) = 0$ only when $x = a_i$ or $x \notin V_i$. Now set $f'(x) = \sum f_{2i}(x)$, a locally finite sum of smooth functions. Then $f' > 0$ on V, and $= 0$ on C_0. Similarly set $f''(x) = \sum f_{2i+1}(x)$. Then $f'' > 0$ on $U - C_1$, and $= 0$ on C_1. Finally define $\phi : U \to I$ by $\phi(x) = \frac{f'(x)}{f'(x)+f''(x)}$. Then $\phi^{-1}(0) = C_0$, $\phi^{-1}(1) = C_1$; further $\nabla\phi = (f''\nabla f' - f'\nabla f'')/(f' + f'')^2 = \sum k_j \nabla f_j$ for suitable C^∞-smooth functions k_j.

Next let $\{W_p\}_{p=1}^{\infty}$ be a countable open covering of $U - (C_0 \cup C_1) = \phi^{-1}(0,1)$ by sets which meet only finitely many V_j. Suppose $x \in W_p$ is a critical point of ϕ, so that $0 = \sum k_j(x)\beta_{jk}(x - a_k)$. Now $k_j(x) \neq 0$ in W_p. Since $\beta_{jj}(x) = 0$ only when $x = a_j$ or $x \notin V_j$, we conclude that either $x = a_j$ or a non-trivial linear combination $\sum \gamma_k(x - a_k) = (\sum \gamma_k)x - \sum \gamma_k a_k$ must be $= 0$. The centres $\{a_k\}$ are linearly independent; hence we find that $\sum \gamma_k \neq 0$, i.e., x belongs to the linear span of

$$\left\{a_k \;\middle|\; \exists\ j \geq k \quad \text{such that} \quad V_j \cap W_p \neq \emptyset\right\} .$$

Let M_p be the intersection of W_p with this linear span; then M_p is a finite-dimensional manifold containing all the critical points of $\phi\big|_{W_p} \to \mathbf{R}^1$. Now apply the Morse-Sard theorem to $\phi\big|_{M_p} \to \mathbf{R}^1$; the set of critical values of this function has μ_1-measure 0. The set of critical values of $\phi : U \to \mathbf{R}^1$ is a subset of a countable union of such null sets, hence itself is a null set of $\mathbf{R}^1$. This completes the proof.

Proof of Theorem 6.1. Let $U \subset \mathcal{H}$ be an open set, $\phi : U \to F$ be a continuous map, and $\varepsilon(\cdot)$ a continuous positive function on U. The object is to show that $\exists$ residual C^∞ $\psi : U \to F \ni \|\phi(x) - \psi(x)\| < \varepsilon(x) \quad \forall x \in U$.

First let $\{U_i\}_{i=1}^{\infty}$ be a countable cover of U by open balls in U with centres at points $\{a_i\}_{i=1}^{\infty}$ such that $\|\phi(x) - \phi(a_i)\| < \varepsilon(x)\ \forall x \in U_i$. Let $\{V_i\}_{i=1}^{\infty}$ be a locally finite scalloped refinement as in Lemma 6.3, and a subordinate partition of unity $\{\lambda_i\}_{i=1}^{\infty}$ using the residual functions in Lemma 6.4. Now set $b_i = \phi(a_i)$, $i \geq 1$, and define the C^∞-map $\psi : U \to F$ by $\psi(x) = \sum\limits_{i=1}^{\infty} \lambda_i(x) b_i$. Then $\|\phi(x) - \psi(x)\| \leq \sum\limits_{i=1}^{\infty} \lambda_i(x)\|\phi(x) - b_i\| < \varepsilon(x)$ $\forall x \in F$.

Now use the proof of Lemma 6.3; this proof yields a countable open cover $\{W_p\}_{p=1}^{\infty}$ of U, each element of which meets only finitely many V_i. If dim span $\{b_j \mid V_j \cap W_p \neq \emptyset\}$ $< \dim F$, then $\psi(W_p)$ lies in a proper linear subspace of F, hence is meager. Otherwise the set of critical points (resp. critical values) of $\psi\big|_{W_p} \to F$ lies in a finite dimensional manifold (resp. linear subspace). The Morse-Sard theorem ensures that this map has a meager set of critical values in F (more precisely, has Lebesgue m_p-measure zero where m_p is the dimension of the linear span of the relevant b_i associated with W_p). Hence the set of critical values of $\psi : U \to F$ is contained in a countable union of meager sets, and is therefore meager, by Baire's theorem. This completes the proof of Theorem 6.1.

§7. Contributions of J. Wells and K. Sundaresan

John Wells in 1968 (cf. [63]) showed that the techniques used in preceding results could not be used in certain spaces. Specifically he showed that in the space c_0 one cannot have a real valued function with bounded support and with a uniformly continuous derivative.

K. Sundaresan (cf. [61]) carried these ideas further, and showed that if a Banach space admits a non-trivial function with bounded support and having a uniformly continuous derivative then the space must be super-reflexive.

The result of Wells ([63]) turns out to be a special case of Sundaresan's theorem. However we shall first present the theorem of Wells, for it has a simple proof which does not require the machinery used by Sundaresan, and also because it showed, for the first time, that in the space c_0, C^2-fine approximation of a C^2-function by a C^∞ function is not possible.

Theorem 7.1. *(cf. [63]). Let $f \in C^1(c_0, \mathbf{R})$ with a uniformly continuous derivative $Df(\cdot)$. Then the support of f is unbounded.*

Proof. Suppose the statement is not true. Then $\exists\ f \in C^1(c_0, \mathbf{R})$ with the properties: $f(0) = 1, f(x) = 0$ for $\|x\| \geq 1$, and Df is uniformly continuous. Let $N \ni \|h\| \leq \frac{1}{N} \Rightarrow \|Df(x+h) - Df(x)\| \leq \frac{1}{2}$. Because of the mean value theorem, we then assert

$$\|h\| \leq \frac{1}{N} \Rightarrow |f(x+h) - f(x) - Df(x) \cdot h| \leq \frac{1}{2}\|h\| \ .$$

Let E be the subset of c_0 consisting of x such that the $2^N - 1$ of the first 2^N components of x have absolute value $\frac{1}{N}$, the remaining component has absolute value less than or

equal to $\frac{1}{N}$, and all components after the first 2^N components are zero. The set E is connected and even, hence we can choose, inductively, $h_1, h_2, \ldots, h_N \in E \ni Df(h_1 + \cdots + h_{k-1}) \cdot h_k = 0$ and $h_1 + \cdots + h_k$ has at least 2^{N-k} components equal to $\frac{k}{N}$.

It follows that $\|h_1 + \cdots + h_N\| = 1$, and

$$\begin{aligned} |f(h_1 + \cdots + h_N) - f(0)| \le & \sum_{k=1}^{N} |f(h_1 + \cdots + h_k) - \\ & f(h_1, \ldots, h_{k-1}) - Df(h_1 + \cdots + h_{k-1}) \cdot h_k| \\ \le & \sum_{k=1}^{N} \frac{1}{2}\|h_k\| = \frac{1}{2} \sum_{k=1}^{N} \frac{1}{N} = \frac{1}{2}, \end{aligned}$$

which is a contradiction. This proves the theorem of Wells.

We shall next turn to the work of Sundaresan ([61]) mentioned above. The main result of his paper requires a good deal of machinery which we shall proceed to explain.

First some definitions and terminology are in order. A Banach space E is said to be *smooth* if for all $x \neq 0$, $x \in E$,

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} = G_x(y)$$

exists $\forall y \in E$. If this limit exists at a point $x \neq 0$, $\forall y \in E$ then it is known that $G_x \in E^*$ the dual of E, and also $\|G_x\| = 1$ (see Kothe [29]). A smooth Banach space E is said to be *uniformly smooth* if the preceding limit is uniform $\forall x, y$ with $\|x\| = 1 = \|y\|$. The homogeneity of the norm implies the following lemmas. If E, F are Banach spaces, a function $f : E \to F$, having a uniformly continuous derivative on a set $A \subset E$ will be said to be *U.C.D. on A*. A Banach space E is said to be U^1-*smooth* if $\exists$ *U.C.D.* real function on E with bounded support.

Lemma 7.2. *A Banach space E is uniformly smooth iff the norm is U.C.D. on regions* $R(\lambda, \mu) = \{x \in E \mid \lambda < \|x\| < \mu \text{ where } 0 < \lambda < \mu\}$.

Lemma 7.3. *The norm in a Banach space E is U.C.D. on regions $R(\lambda, \mu)$ iff the norm is uniformly differentiable on bounded sets away from the origin.*

We shall also need the following lemmas, which are either results on differential calculus in Banach spaces (see Dieudonné [9], Lang [32]), or consequences of the preceding definitions.

Lemma 7.4. *Let E be a Banach space, $f : E \to \mathbf{R}$ a U.C.D. function, and Df the derivative of f. Then*

(a) if U is a bounded subset in E, then $f\big|_U$ is Lipschitzian: $EM > 0 \ni \forall x, y \in U$, $\|f(x) - f(y)\| \leq M\|x - y\|$;

(b) if the support of f is bounded then f is Lipschitz everywhere in E, in particular f is uniformly continuous.

Lemma 7.5. *Suppose E is a U^1-smooth Banach space, and $\lambda > 0$, then $\exists$U.C.D. real function f on $E \ni f(0) = 1$, and $f(x) = 0$ for $\|x\| \geq \lambda$.*

Lemma 7.6. *If f and g are two U.C.D. real functions on a Banach space E and the support of f (or the support of g) is bounded, then fg is U.C.D. with bounded support.*

Lemma 7.7. *If E, F, G are three Banach spaces, $f : E \to F$, $g : F \to G$ are U.C.D. functions such that the derivatives Df, Dg are bounded on $E \to L(E, F)$, and on $F \to L(F, G)$, respectively, then the composite $g \circ f$ is U.C.D.*

For the next lemma we refer ther reader to Nemirovski and Semenov ([45]).

Lemma 7.8. *Suppose E is a uniformly convex and uniformly smooth Banach space. Then the restrictions of the uniformly continuously differentiable functions on E to any closed ball $\overline{U_r(0)}$ is dense in the space of uniformly continuous functions on $\overline{U_r(0)}$ with the uniform topology.*

We now have to turn to the concepts of super-reflexive Banach spaces and ultra-powers of normed linear spaces. If E, F are Banach space, then E is said to be *finitely represented* in F, in symbols $E \ll F$, if for each finite dimensional subspace of E and positive number ε, $\exists$ subspace Y of F, depending on X and ε, $\ni$ there is an isomorphism $T : X \overset{\text{onto}}{\to} Y$ with $\|T\|\|T^{-1}\| \leq 1 + \varepsilon$. A Banach space F is said to be *superreflexive* if $E \ll F$ implies E is reflexive. For basic results on this topic see James [24], and shall be content with stating one known fact which we shall use.

Fact 7.9. (cf. Enflo [11]). A Banach space E is super reflexive $\Longleftrightarrow$ E is isomorphic with a uniformly smooth Banach space.

Let S be an infinite set and Γ a non-trivial (free) ultrafilter on S. If f is a bounded real-valued function on S let $\lim\limits_{\Gamma} f(s) = \sup\left[\lambda \middle| \{t \in S : f(t) > \lambda\} \in \Gamma\right]$. If $(E, \|\ \ \|)$ is a normed linear space, and f is a bounded E-valued function on S, then let $|f| = \lim\limits_{\Gamma} \|f(s)\|$. Then $|\ \ |$ is a semi-norm on the vector space V of bounded E-valued functions on S, and the quotient space of V modulo the kernel of $|\ \ |$ equipped with the quotient norm is known as the *ultrapower* of E *associated with the pair* S, Γ, and is denoted by $E(S, \Gamma)$. The following facts are known (cf. [61]).

Fact 7.10. If E is a Banach space then $E(S, \Gamma)$ is a Banach space.

Fact 7.11. If E, F are Banach spaces then $E \ll F \Leftrightarrow E$ is isometric with a subspace

of an ultrapower $F(S,\Gamma)$ of F.

Before turning to the next lemma, the following remarks are in order.

Remark 1. If E is a uniformly smooth Banach space, so that the norm of E is $U.C.D.$ on regions $R(\lambda,\mu)$ then by composing the norm of E with a suitable C^1-function $\mathbf{R}^1 \to \mathbf{R}^1$ and using Lemma 7.7, we find that E is U^1-smooth. Then apply Lemmas 7.5, 7.7, and use the fact that the norm of E is $U.C.D.$ on $R(\lambda,\mu)$, and we verify that if E is uniformly smooth, and r,ε are two positive numbers, then $\exists$ $U.C.D.$ function $f : E \to \mathbf{R}^1$ $\ni 0 \le f \le 1$, $f \equiv 1$ on $B_r(0)$, and $f \equiv 0$ outside $\overline{B_{r+\varepsilon}(0)}$. The next lemma uses these remarks and the lemma of Nemirovski and Semenov (cf. [45]).

The support of a real function is denoted by $\text{supp}(f)$.

Lemma 7.12. *(Sundaresan [61]) Suppose G is a nonempty open subset of a super-reflexive space E. Then $\exists U.C.D.$ function $f : E \to \mathbf{R}^1$, $\ni 0 \le f \le 1$ and* $\text{supp}(f) = G$.

Proof of Lemma. We shall assume w.l.g that E is uniformly convex and uniformly smooth. First suppose G is a bounded open set, and let $C = E - G$. Let $g : E \to \mathbf{R}^1$ be defined by: $g(x) = d(x,C) =$ distance of x from C. Let $r > 0 \ni G \subset \bar{G} \subset B_{r/2}(0)$. Consider the restriction of the uniformly continuous function g to $\overline{B_r(0)}$.

By the Lemma of Nemirovskii and Semenov, $\exists U.C.D.$ functions f_n on $E \to \mathbf{R}^1 \ni$ $\sup_{x\in B_r(0)} |f_n(x) - f(x)| < \frac{1}{n}$, $n = 1,2,\ldots$. Now use the preceding remarks; we find that $\exists U.C.D.$ function $\phi : E \to \mathbf{R}^1$, $\ni \phi \equiv 1$ on $\overline{B_{r/2}(0)}$, $\phi \equiv 0$ outside $B_r(0)$; let $h_n = \phi f_n$. Let $\alpha_n(\cdot)$ be the C^1-functions on $\mathbf{R}^1 \to \mathbf{R}^1 \ni \text{supp}(\alpha_n) = (\frac{1}{n},\infty)$, $\alpha_n(t) = 0$ if $t \le \frac{1}{n}$, $\alpha_n(t) = 1$ if $t \in [\frac{2}{n},\infty)$ and $0 \le \alpha_n \le 1$ for $n = 1,2,\ldots$. Let $g_n = \alpha_n(h_n)$, $g = \sum_{n=1}^{\infty} \frac{1}{2^n} g_n$. Then g is a $U.C.D.$ function $E \to \mathbf{R}^1$, $\text{supp}(g) = G$, and $0 \le g \le 1$. This proves the

lemma under this additional assumption.

Now suppose G is an arbitrary open set, then we write $G = \bigcup_{n=1}^{\infty} G_n$, where $G_n = B_n(0) \cap G$, $n = 1, 2, \ldots$. Each G_n is a bounded set, hence by the preceding paragraph, $\exists$ $U.C.D.$ function $f_n : E \to \mathbf{R}^1 \ni \operatorname{supp}(f_n) = G_n$, and $0 \le f_n \le 1$. Let $f = \sum_{n=1}^{\infty} \frac{1}{2^n} f_n$. Then this f has the required properties.

The next two theorems show that U^1-smoothness is finitely inherited.

Theorem 7.13. *(K.S. [61]). If E is U^1-smooth then every ultrapower $E(S,\Gamma)$ is U^1-smooth.*

Proof of Theorem. We shall denote the norms in $E, E(S,\Gamma)$ by $\|\ \ \|$, and $|||\ \ |||$ respectively. E is assumed U^1-smooth, hence $\exists U.C.D.$ function $f : E \to \mathbf{R}^1 \ni f \not\equiv 0$, with $\operatorname{supp}(f) \subset \overline{B_1(0)}$. Let $\tilde{x} \in E(S,\Gamma)$ and let $\{x(s)\}_{s\in S}$ be a representive of $\tilde{x}$. f is a bounded function, hence $\lim_{\Gamma} f(x(s))$ exists. Let $\{y(s)\}_{s\in S}$ be another representative of $\tilde{x}$. f is uniformly continuous by Lemma . Hence if $\varepsilon > 0$ $\exists \delta > 0 \ni \|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Since $\{x(s)\}_{s\in S}$, $\{y(s)\}_{s\in S}$, represent the same equivalence class $\tilde{x} \in E(S,\Gamma)$, $\exists$ $J_\delta \in \Gamma \ni s \in J_\delta \Rightarrow \|x(s) - y(s)\| < \delta$. Hence $\forall s \in J_\delta$, $|f(x(s)) - f(y(s))| < \varepsilon$. Hence if $f^*(\tilde{x}) = \lim_{\Gamma} f(x(s))$, then f^* is a real-valued function on $E(S,\Gamma)$. The support of f is in $\overline{B_1(0)} \subset E$, hence

$$f^*(\tilde{x}) \neq 0 \Rightarrow \exists J \in \Gamma \ni \|x(s)\| \le 1 \quad \forall s \in J \ ,$$

hence $|||\tilde{x}||| \le 1$. Hence $\operatorname{supp}(f^*)$ is contained in the unit ball of $E(S,\Gamma)$.

Now $Df : E \to E^*$ is a uniformly continuous mapping with bounded range; hence proceeding as in the preceding paragraph we verify that if $\tilde{x}, \tilde{y} \in E(S,\Gamma)$ and $\{x(s)\}_{s\in S}$,

$\{y(s)\}_{s\in S}$ are representatives of $\tilde{x},\tilde{y}$ respectively, then $\lim_{\Gamma} Df(x(s))(y(s))$ is independent of the representatives $\{x(s)\}$, $\{y(s)\}$ of $\tilde{x},\tilde{y}$ respectively. Now defined $\ell_{\tilde{x}}(\tilde{y}) = \lim_{\Gamma} Df(x(s))(y(s))$, $\tilde{x},\tilde{y} \in E(S,\Gamma)$. We then verify that $\ell_{\tilde{x}}$ is a continuous linear functional on $E(S,\Gamma)$, since Df is bounded.

Now if $\{h(s)\}_{s\in S} \in \tilde{h}$ then

$$f^*(\tilde{x}+\tilde{h}) = \lim_{\Gamma} f(x(s)+h(s)) = \lim_{\Gamma}\Big\{ f(x(s)) + Df(x(s))(h(s)) + \theta_{x(s)}(h(s))\Big\} ,$$

where we note that since f is a *U.C.D.* function, given $\varepsilon > 0$, $\exists\delta > 0 \ni |\theta_x(y)| \le \varepsilon\|y\|$ if $\|y\| \le \delta\ \forall x \in E$. Then $\exists$ set $J \in \Gamma \ni \forall s \in J$, $|\theta_{x(s)}(h(s))| \le \varepsilon\|h(s)\|$. Hence $\lim_\Gamma |\theta_{x(s)}(h(s))| \le \varepsilon\ |\!|\!|\tilde{h}|\!|\!|$ if $|\!|\!|\tilde{h}|\!|\!| \le \delta$ and f^* is differentiable at $\tilde{x}$ with $Df^*(\tilde{x}) = \ell_{\tilde{x}}$.

Df is uniformly continuous on $E \to E^*$, hence we verify that the map $Df^* : E(S,\Gamma) \to (E(S,\Gamma))^*$ is uniformly continuous, once again working with suitable members of Γ as previously. Thus $E(S,\Gamma)$ is U^1-smooth. This completes the proof.

As corollaries we obtain the next results.

Corollary 7.14. *If a Banach space E is U^1-smooth and $F \ll E$, then F is U^1-smooth.*

Proof. $F \ll E$ if and only if F is isometric with a subspace of some ultrapower $E(S,\Gamma)$ of E. The corollary then follows.

Corollary 7.15. *If a Banach space E is superreflexive, then it is U^1-smooth.*

Proof. A superreflexive Banach space is isomorphic with a uniformly smooth Banach space. Also U^1-smoothness is invariant, under isomorphisms. The corollary follows.

The second theorem of Sundaresan in this context is as follows.

Theorem 7.16. *(cf. [61]). If E is a U^1-smooth Banach space, then it is reflexive.*

Proof. Let $0 < \theta < 1$. By Lemma $\exists U.C.D.$ real-valued function f on $E \ni f(0) = 1, f(x) = 0$ if $\|x\| \geq \theta/4$. Since f is $U.C.D.$, therefore if $0 < \varepsilon < 1$, $\exists$ positive integer $M \ni$ if $h \in E, \|h\| \leq \frac{1}{m}$ then

$$(A) \qquad \left|f(x+h) - f(x) - Df(x)h\right| \leq \varepsilon\|h\| .$$

Suppose E is non-reflexive. Then by a theorem of James (cf. [24]) it follows that $\exists$ set X containing the set of positive integers and a subspace L of the Banach space $B(X)$ of bounded real-valued functions on X with the supremum norm, isometric with E, admitting a sequence $\{z_n\}_{n\geq 1}$, $\ni$ for $n \geq 1$

$$\begin{aligned} z_n(i) &= \theta \qquad 1 \leq i \leq n, \quad i \in W, \\ z_n(i) &= 0 \qquad i > n, \quad i \in W , \end{aligned}$$

and

$$|z_n(t)| \leq 1 \qquad \text{for} \qquad t \in X - W .$$

Let $x_{n,0} = \frac{1}{2}z_n, x_{0,n} = -\frac{1}{4}z_n$ for $n \geq 1$, and $x_{n,k} = \frac{3}{4}z_n - \frac{1}{4}z_{n-k}$ if $n \geq 1$, $k \geq 1$. Then $\|x_{n,k}\| \leq 1$ for all pairs of integers (n,k) for which $x_{n,k}$ is defined. Now define the polynomial path P in L by

$$P = \bigcup_{i=0}^{2^M-1} \left[x_{2^M-i,i}, x_{2^M-i-1,i+1}\right]$$

where M is the positive integer chosen to satisfy (A) in the preceding paragraph. Consider the derivative $Df(0)$. By our choice of $x_{n,k}$, $Df(0)(x_{2^M,0}) = 0$ if and only if

$Df(0)(x_{0,2^M}) = 0$, and $Df(0)(x_{2^m,0} \gtrless 0 \Leftrightarrow Df(0)\,(x_{0,2M}) \lessgtr 0$. The path P is connected, hence $\exists \xi \in P \ni Df(0)(\xi) = 0$. If

$$\xi \in \left[x_{2^M - i_0, i_0}, x_{2^M - i_0 - 1, i_0 + 1}\right] ,$$

then

$$\begin{aligned} \xi(j) &= \frac{1}{2}\theta \quad \text{if} \quad 1 \le j \le 2^M - i_0 - 1, \quad j \in W , \\ \xi(j) &= -\frac{1}{4}\theta \quad \text{if} \quad 2^M - i_0 + 1 \le j \le 2^M, \quad j \in W , \\ \xi(j) &\in \left[-\frac{1}{4}\theta, \frac{1}{2}\theta\right] \quad \text{if} \quad j = 2^M - i_0, \quad j \in W , \end{aligned}$$

and

$$\|\xi\| \le 1 .$$

Thus if $\xi(j_0) < \frac{1}{2}\theta$ for some $j_0 \in W$, $1 \le j \le 2^M$ (which is the case if $\xi(j_0) \in \left(-\frac{\theta}{4}, \frac{\theta}{2}\right)$ or $\xi(j_0) = -\frac{\theta}{4}$), then $\xi(j) = -\frac{1}{4}\theta\ \forall j \in W$, $j_0 + 1 \le j \le 2^M$. Now if $2^M - i_0 - 1 \ge 2^{M-1}$, choose $\xi_1 = \xi/M$, otherwise $\xi_1 = -\xi/M$. The ξ_1 thus chosen has the properties: $\|\xi_1\| \le \frac{1}{M}$, $Df(0)(\xi_1) = 0$, and $\xi_1(j) \ge \frac{\theta}{4M}$ for at least 2^{M-1} values of $j \in W$, $1 \le j \le 2^M$.

Now consider $Df(\xi_1)$. As before $\exists \xi' \in P \ni Df(\xi_1)(\xi') = 0$. From the properties of ξ noted in the preceding paragraph, since $\xi_1 = \pm\frac{\xi}{M}$, the restriction of ξ_1 to the set $Q = \{j \mid 1 \le j \le 2^M\} \subset W$, has range either in the set $\{\frac{\theta}{2M}, -\frac{\theta}{4M}\}$ or $\{-\frac{\theta}{2M}, \frac{\theta}{4M}\}$ except possibly for one value of $j \in Q$. These observations imply either

(i) $(\xi_1 + \frac{\xi'}{M})(j) \ge \frac{2\theta}{2M}$, or

(ii) $(\xi_1 - \frac{\xi'}{M})(j) \ge \frac{2\theta}{2M}$, for at least 2^{M-2} integers $j \in Q$. Let $\xi_2 = \frac{\xi'}{M}$ or $= -\frac{\xi'}{M}$ according as (i) or (ii) is the case. Repeating this procedure inductively it follows that $\exists$ sequence $\{\xi_i\}_{i=1}^M$ in $L \ni \|\xi_i\| \le \frac{1}{4M}$, $Df\left(\sum_{i=1}^{k-1} \xi_i\right)(\xi_k) = 0$, $\sum_{i=1}^{k} \xi_i(j) \ge \frac{k\theta}{4M}$

for $1 \leq k \leq M$. From our choice of $f, M, \{\xi_i\}_{i=1}^{M}, \varepsilon$, together with the inequality $\|\sum_{i=1}^{k} \xi_i\| \geq \frac{k\theta}{4M}$ it follows that

$$1 + \left|f(\sum_{i=1}^{M} \xi_i) - f(0)\right| \quad \sum_{k=1}^{M} \left|f(\sum_{i=1}^{k} \xi_i) - f(\sum_{i=1}^{k-1} \xi_i) - Df(\sum_{i=1}^{k-1} \xi_i)(\xi_k)\right|$$
$$\leq \sum_{i=1}^{M} \varepsilon\|\xi_i\| \leq \varepsilon < 1 \ ,$$

a contradiction; thus the proof is completed.

The next theorem provides a characterisation of U^1-smooth Banach spaces.

Theorem 7.17. *(cf. K.S. [61]). A Banach space E is superreflexive if and only if E is U^1-smooth.*

Proof. From Corollary 7.15 and the Theorem 7.16 it follows that if E is U^1-smooth, and $F \ll E$, then F is reflexive. Thus E is superreflexive. The converse follows from Remark

The Banach spaces $c_0, C(K)$ where K is an infinite compact Hausdorff space are not superreflexive (in fact not even reflexive), the earlier theorem of J. Wells now follows as a consequence of the preceding characterisation.

Corollary 7.18. *If $E = c_0$, or $C(K)$ with K as above, then E is not U^1-smooth.*

§8. Theorems of Desolneux-Moulis

The work of Nicole Desolneux-Moulis was a difficult and major step in obtaining information concerning strong approximation in infinite-dimensional Banach spaces, at a time when such information was extremely scanty. It was realised that the tools that that were known to be effective in finite-dimensional spaces, such as smoothing by convolution, or use of smooth partitions of unity, could not be directly utilised in infinite dimensional spaces., the reasons being rather obvious, viz. partly the lack of a convenient theory of integration comparable to the Lebesgue theory of integration in finite dimensional spaces. The tools devised in [8] provided some of the ideas for the further work which this author was able to carry on (§9). In this section we shall briefly describe some of the theorems proved in [8], with a sketch of some of the arguments.

First we shall describe the theorem regarding C^1-fine approximation established in [8] for the spaces ℓ_p (p being an integer ≥ 2) and for c_0. Earlier in §2 of this chapter we found that the space ℓ_{2p} ($p \geq 1$ an integer) has a C^∞-norm away from $0, c_0$ *allows* an equivalent C^∞-norm, and the space l_{2p+1} ($p \geq 1$ an integer) has a C^{2p}-norm away from 0. In the case of c_0 it is understood in the paper [8] that the original norm is replaced by the equivalent C^∞-norm. The methods used in this paper were the same for the following spaces: ℓ_p ($p \geq 2$ an integer), or the space c_0 (with the equivalent C^∞-norm); it is therefore convenient to use the symbol E^α to denote any one of these spaces, with the C^α-norm ($\alpha = \infty$ for ℓ_{2p}, or c_0, $\alpha = 2p$ for ℓ_{2p+1}).

Theorem 8.1. *([8] Theorem 1). Let Ω be an open set in E^α, F a Banach space. The space $C^\infty(\Omega, F)$ is dense in $C^1(\Omega, F)$ in the C^1-fine topology.*

For the proof of this theorem, suppose $\varepsilon(\cdot)$ is a positive continuous function on Ω and $f : \Omega \to F$ a C^1-mapping. The object is then to construct a C^∞-mapping $g \ni \|g(x) - f(x)\| < \varepsilon(x)$ and $\|Dg(x) - Df(x)\| < \varepsilon(x)\ \forall x \in \Omega$. The following lemmas are used in the proof.

Lemma 8.2. *(Lemma 1, page 295 in [8]) Suppose B_1 is an open ball with centre 0 in E^α, and $f : B_1 \to F$ a C^1-mapping. Suppose $B_0 \subset B_1$ is an concentric open ball and $\eta > 0 \ni \sup_{x\in B_0} \|D^1 f(x)\| < \eta$. Let $\varepsilon > 0$. Then $\exists$ constants λ_0, λ_1 and $\exists g : B_1 \to F \ni g$ is C^α and satisfies:*

$$\sup_{x\in B_0} \|g(x) - f(x)\| < \lambda_0\varepsilon; \quad \text{and} \quad \sup_{x\in B_0} \|Dg(x)\| < \lambda_1\eta\ .$$

The proof of this lemma is carried out in two steps. Let $\{e_n\}_{n=1}^\infty$ be the canonical basis in E^α, i.e., $e_n = \{x_p^n\} \ni x_p^n = 0$ if $p \neq n$, $x_n^n = 1$. Let E_n be the subspace generated by e_p, $1 \le p \le n$, and $E^\infty = \bigcup_{n=1}^\infty E_n$.

Step 1. A function $\bar f : E^\infty \to F$ is constructed $\ni \bar f$ restricted to E_n for each $n = 1, 2, \ldots$ is C^∞ and $\bar f$ is "close" to f.

Step 2. A mapping $\psi(x) : E^\alpha \to E^\infty$ is constructed $\ni \|x - \psi(x)\| < r$ for a suitable number r. Then $g(\cdot)$ is defined by: $g(x) = \bar f\big(\psi(x)\big)$.

A second lemma that is needed is the following.

Lemma 8.3. *(Lemma 2, p. 301 in [8]). For each $x \in \Omega\ \exists$ open ball $B_{r(x)}(x) \subset \Omega$ satisfying*

(a) $\sum_{y,y'\in B_{r(x)}(x)} |\varepsilon(y) - \varepsilon(y')| < \inf_{y\in B_{r(x)}(x)} \frac{\varepsilon(y)}{2}$;

(b) $\sup_{y,y'\in B_{r(x)}(x)} \|Df(y) - Df(y')\| < \inf_{y\in B_{r(x)}(x)} \frac{\varepsilon(y)}{2}$;

(c) *if* $B_{r(x)}(x) \cap B_{r(y)}(y) \neq \phi$ *then* $\frac{1}{4} < \frac{r(x)}{r(y)} < 4$;

(d) $\sup\limits_{x\in\Omega} r(x) < 2$.

The separability of the space implies that there is a countable collection of points $\{a_1, a_2, \ldots\}$ in $\Omega \in \Omega = \bigcup\limits_{n=1}^{\infty} B_{\frac{r(a_n)}{2}}(a_n)$.

The second theorem on strong approximation in [8] is the following.

Theorem 8.4. *([8] Theorem 2, p. 306). Suppose E is a separable infinite dimensional Hilbert space, $\Omega \subset E$ an open set, and F a Banach space. Then the class $C^\infty(\Omega, F)$ is dense in the class $C^{2k-1}(\Omega, F)$ endowed with the C^k fine topology (denoted by $C^k_{2k-1}(\Omega, F)$.*

The proof of Theorem 8.4 proceeds along the lines of proof of Theorem 8.1. One lemma needed in the proof is the following.

Lemma 8.5. *Suppose $F : E \to F$ is a C^{2k-1}-mapping. Suppose B_0 is an open ball with centre 0 contained in E and $\eta > 0$ a number $\ni \sup\limits_{x\in B_0} \|D^{2k-1}f(x)\| < \eta$. Let $\varepsilon > 0$ be arbitrary, and set $r = \varepsilon/\eta$. Then $\exists (k+1)$ constants $\lambda_i (0 \leq i \leq k)$ and $\exists$ C^∞ function g satisfying: for $0 \leq i \leq k-1$, $\sup\limits_{x\in B_0} \|D^i g(x) - D^i f(x)\| < \lambda_i \varepsilon r^{k-1-i}$, and $\sup\limits_{x\in B_0} \|D^k g(x)\| < \lambda_k \eta$.*

As in the proof of Lemma 8.2 for the proof of Lemma 8.5 one constructs a function $\bar{f}$: $E^\infty \to F \ni \bar{f}$ restricted to E_n for $n = 1, 2, \ldots,$ is C^∞ and further $\bar{f}$ is "close" to f.

The second step in the proof consists in exhibiting a function $\psi : E \to E^\infty \ni$

(a) ψ is C^∞,

(b) for each $x \in E$ $\exists$ positive integer $n(x)$ and $\exists$ neighbourhood U_x of $x \in E \ni$ $\psi(U_x) \subset E_{n(x)}$;

(c) $\exists$ constants $C_1 \ni \|D\psi(x)\| < C_1$;

(d) $\exists$ constants $C_i \ni \|D^i\psi(x)\| < C_i r^{i-i}$.

Then g is defined by

$$g(x) = \bar{f}(\psi(x)) + \sum_{i \leq i \leq k-1} \frac{1}{i!} L^i(\psi(x)) \cdot (x - \psi(x))^{(i)}$$

for suitable functions L^i.

The proof of Theorem 8.5 is completed as follows. Let $\varepsilon(\cdot)$ be a positive continuous function on $\Omega, f \in C^{2k-1}(\Omega, F)$. Then the objective is to exhibit a function $g \in C^\infty(\Omega, F) \ni$ for $0 \leq \alpha \leq k$ $\|D^\alpha g(x) - D^\alpha f(x)\| < \varepsilon(x) \forall x \in \Omega$. For this purpose another crucial lemma which is needed is the following.

Lemma 8.6. *For any point $x \in \Omega, \exists$ open ball $B_{\rho(x)}(x) \subset \Omega$ satisfying:*

(a) $\sup_{y,y' \in B_{\rho(x)}(x)} |\varepsilon(y) - \varepsilon(y')| < \inf_{y \in B_{\rho(x)}(x)} \frac{\varepsilon(y)}{2}$;

(b) $\sup_{y,y' \in B_{\rho(x)}(x)} \|D^k f(y) - D^k f(y')\| < \inf_{y \in B_{\rho(x)}(x)} \varepsilon(y)$,

(c) if $B_{\rho(x)}(x) \cap B_{\rho(y)}(y) \neq \phi$, *then* $\frac{1}{4} < \frac{\rho(x)}{\rho(y)} < 4$;

(d) $\sup_{x \in \Omega(x)} \rho(x) < 2$.

Theorems 8.1 and 8.5 are the two fundamental theorems on strong approximation established in [8]. To describe the third interesting theorem (which is a consequence of Theorem 8.1) in [8] we should first state a definition.

Let Ω be an open set in a separable infinite dimensional Hilbert space E, F a separable Hilbert space of finite or infinite-dimension. A mapping $f : \Omega \to F$ is said to be of *Sard type* if the set of its critical values has no interior points.

Theorem 8.7. *(cf. [8], p. 331) The class of C^∞ mappings $\Omega \to F$ which are of Sard type is dense in $C^1(\Omega, F)$ endowed with the C^1 fine topology.*

We shall give a brief sketch of the proof. To be specific, the function $g(x)$ exhibited in Theorem 8.1 of this section and satisfying

$$\left.\begin{array}{l}\|g(x)-f(x)\|<\lambda_0\varepsilon(x),\\ \\ \|Dg(x)-Df(x\|<\lambda_1\varepsilon(x)\end{array}\right\}\quad Ax\in\Omega\ ,$$

λ_0, λ_1 being fixed constants, is itself a function of Sard type. In the proof of Theorem 8.1, Ω was shown to be a countable union: $\Omega = \bigcup\limits_{n=1}^{\infty} B_{r_n/2}(a_n)$ (where we have written r_n for the earlier $r(a_n)$, $n = 1, 2, \ldots$). Let $x \in \Omega$, and n the least integer $m \ni x \in B_{r_m/2}(a_m)$. On $B_{r_n/2}(a_n)$ the function $g(\cdot)$ coincides with the function g_n where

$$\begin{aligned} g_n(x) &= \sum_{1\le p\le n} \mu_p(x)\big[f(a_p)+Df(a_p)\cdot(x-a_p)+\delta_p(x)\big]\ , \\ \mu_p(x) &= \phi_p(x)\big(1-\phi_{p-1}(x)\big)\cdots\big(1-\phi_1(x)\big)\ , \end{aligned}$$

and $\delta_p(x)$ is constructed through the following steps:

Let $\psi_n(\cdot) : E \to E^\infty$ be the differentiable function $\ni \|y - \psi_n(y)\| < \frac{r_n}{2^n}$.

There is an open ball $\Omega'(x)$ with centre x, contained in $B_{r_n/2}(a_n)$ and an integer $m \ni \psi_n(\Omega'(x)) \subset E_m$ (E_m of finite dimension m); for any integer $p (1 \le p \le n)$, $\exists\, C^\infty$ mapping $\bar{\delta}_p : E_m \to F \ni \delta_p(\cdot) = \bar{\delta}_p(\psi_n(\cdot))$.

Let $\Omega(x)$ be an open ball centre $x \ni \Omega(x) \subset \overline{\Omega(x)} \subset \Omega'(x) \subset B_{r_n/2}(a_n)$.

The open balls $\Omega(x), x \in \Omega$, form an open covering of Ω, hence $\Omega = \bigcup\limits_{i=1}^{\infty} \Omega(x_i)$, for a countable collection $\{x_1, x_2, \ldots\}$ of points in Ω. We shall write Ω_i as short for $\Omega(x_i), i = 1, 2, \ldots$. Then *exists* integer $n_i \ni$ in Ω'_i:

$$g(x) = \sum_{p\le n_i} \mu_p(x)\Big[f(a_p)+Df(a_p)\cdot(x-a_p)+\bar{\delta}_p\big(\psi_{n_i}(x)\big)\ .$$

The remaining part of the proof of Theorem 8.7 consists in showing that the complement of the set of critical values of g on Ω'_i is a countable intersection of open dense sets, F is a Banach, hence a Baire space, and it follows that g is of Sard type.

We shall now simply state two more theorems established in [8].

Suppose E^α is as in Theorem 8.1 (i.e., E^α is one of the spaces $\ell_p, p \geq 1$ an integer, c_0, with a C^α-norm), M a paracompact manifold modelled on E^α, and N a paracompact manifold modelled on the Banach space F. For any integer $j \geq 0$ denote by $C^j(M,N)$ the class of C^j-mappings $M \to N$.

Theorem 8.8. *(cf. [8], p. 325) If M and N are C^α-manifolds, then $C^\infty(M,N)$ is dense in $C^1(M,N)$ endowed with the C^1-fine topology.*

Now suppose E is a separable infinite dimensional Hilbert space, M a C^∞-manifold modelled on E, and N a paracompact manifold modelled on F.

Theorem 8.9. *(cf. [8], p. 328) The class $C^\infty(M,N)$ is dense in $C^{2k-1}(M,N)$ endowed with the C^k fine topology.*

§9. C^k-fine approximation of C^k by C^∞–: a theorem of Heble

I. We shall give a proof of the following theorem (cf. [17], [18], [19]) in this section:

Theorem 9.1. *Let Ω be a nonempty open set in a separable real Hilbert space $\mathcal{H}$, F a real Banach space, $f : \Omega \to F$ a C^k-smooth mapping $k \geq 0$ being a given integer, differentiability being always understood in the Fréchet sense, and $\varepsilon(\cdot)$ a continuous positive function on Ω. Then $\exists g : \Omega \to F \ni g$ is C^∞-smooth and satisfies $\forall$ integers $j \in [0, k]$, $\|D^j g(x) - D^j f(x)\|_j < \varepsilon(x)$ $\forall x \in \Omega$.*

This means: $C^\infty(\Omega, F)$ is dense in $C^k(\Omega, F)$ in the C^k fine topology (see Chapter II, §2, for the definition of the C^k fine topology).

Remark. This theorem had been already proved earlier in the special case $k = 0$ (cf. [5], [30]), the proof here (reproduced from [18, [19]) is for any integer $k \geq 0$. This theorem consists of 2 parts, both of which are proved below:

A: *$\exists$ dense open subset $\Omega' \subset \Omega$, and $\exists \bar{g} \in C^k(\Omega, F) \ni \bar{g} \in C^\infty(\Omega', F)$ and satisfies:*

$$\forall \text{ integers } j \in [0, k], \qquad \|D^j \bar{g}(x) - D^j f(x)\|_j < \varepsilon(x) \quad \forall x \in \Omega \ .$$

Before stating part B of Theorem 1, we should explain some notation. Let Ω' be the particular dense open subset of Ω and $\bar{g}$ the special mapping in $C^k(\Omega, F)$ exhibited in Theorem A. We then define for $x \in \Omega$,

$$\bar{\varepsilon}(x) = \max_{0 \leq j \leq k} \|D^j \bar{g}(x) - D^j f(x)\|_j, \qquad \text{and} \quad \sigma(x) = \frac{1}{2}\big(\varepsilon(x) - \bar{\varepsilon}(x)\big) \ .$$

B: $\exists g \in C^\infty(\Omega, F) \ni \forall$ *integers* $j \in [0, k], \|D^j g(x) - D^j \bar{g}(x)\|_j \quad < \sigma(x)$.

The proofs of parts A and B together will complete the proof of Theorem 9.2. We shall now first prove Theorem A. It is necessary to explain some notation which will be consistently used and also explain some techniques, such as e.g. a Leibnitz formula, and localising - a word which we shall use to mean multiplying by a suitable C^∞ function $\phi : \mathcal{H} \to [0,1]$ with a bounded support.

II. (a) Notation. Let $\Omega, F, f, \varepsilon$ be as stated in the theorem. The norm in the Hilbert space will be denoted by $\|\ \ \|_{\mathcal{H}}$, and the norm in the Banach space F by $\|\ \ \|$. The subscript j after a norm shall denote that the norm is that of a j-linear mapping, e.g. $\|D^j f(x)\|_j$, where $D^j f(x) : \underbrace{\mathcal{H} \times \cdots \times \mathcal{H}}_{j} \to F$; or $\|D^j \phi(x)\|_j : \underbrace{\mathcal{H} \times \cdots \times \mathcal{H}}_{j} \to \mathbf{R}^1$. $B_r(z)$ shall denote, as usual the open ball: $[x \,\big|\, \|x - z\|_{\mathcal{H}} < r]$, or $[x \,\big|\, \|x - z\| < r]$, etc. in the particular Banach space under reference. The closure of a set E is denoted (as is customary) by $\bar{E}$. The unit interval $[0 \le x \le 1]$ in $\mathbf{R}^1$ is denoted by $[0,1]$. ϕ shall always denote the particular C^∞ function: $H \to [0,1]$ explained in subsection II (c) below of this section, $\phi_\rho(x - c)$ denotes the same function after scaling (by $\frac{1}{\rho}$) and a translation of the origin by c. Capital letters C, K, M will denote positive real constants, often depending on a positive integer and one or more subscripts, e.g. $C(n), K_n$ or $M_{j,i}(n)$.

(b) An extended Leibnitz formula. Suppose $\phi : H \to \mathbf{R}^1$ and $g : \Omega \to F$, both being at least C^k-smooth. Suppose j is an integer $\ni 0 \le j \le k$. Then a product such as $\phi g : H \to \mathbf{R}^1$, is also j-times differentiable for $0 \le j \le k$ and we have

$$D^j\big(\phi(x)g(x)\big) = D^j\phi(x)\cdot g(x) + \binom{j}{1} D^{j-1}\phi(x)\cdot Dg(x) + \cdots + \phi(x)\cdot D^j g(x)$$

where the products on the rightside are tensor products, the coefficients $\binom{j}{i}$ being the

binomial coefficients: $\binom{j}{i} = \frac{j!}{i!(j-1)!}$, $0 \le i \le j$. It follows that

$$\|D^j(\phi(x)g(x))\|_j \le \|D^j\phi(x))\|_j \; \|g(x)\| + \binom{j}{1}\|D^{j-1}\phi(x)\| \; \|Dg(x)\| + \cdots + \phi(x)\|D^j g(x)\|_j.$$

(c) Localising. As mentioned above, we shall use this term for multiplication by a particular C^∞-function $\phi : H \to [0,1]$ with bounded support. This function ϕ is defined as follows. Let

$$\alpha(t) = \begin{cases} e^{-\frac{1}{(t-a)(b-t)}}, & t \in (a,b), \quad 0 < a < b \le 1 \; ; \\ 0, & t \notin (a,b); \end{cases}$$

$$\beta(x) = \frac{\int_x^b \alpha(t)dt}{\int_a^b \alpha(t)dt}, \qquad x \in \mathbf{R}^1 \; .$$

For our purposes we shall always choose $a = \frac{1}{2}, b = 1$, and define ϕ by $\phi(x) = \beta\left(\|x\|^2_{\mathcal{H}}\right)$, $x \in \mathcal{H}$. We shall have occasion to use the modified functions $\phi_\rho(x) = \phi(\frac{1}{\rho}x)$, $0 < \rho < 1$, and also $\phi_\rho(x-c) = \phi\left(\frac{1}{\rho}(x-c)\right)$, c being a fixed vector in H.

We shall need precise estimates for $\|D^j\phi(x)\|_j, j = 0, 1, \ldots, k$, as also $\|D^j\Phi^{\rho_1,\rho_2\ldots\rho_n}(x)\|_j$ where $\Phi^{\rho_1\ldots\rho_n}(x) \stackrel{def}{=} \phi_{\rho_1}(x-x_1)\phi_{\rho_2}(x-x_2)\ldots\phi_{\rho_n}(x-x_n)$; here $\{\rho_n\}_{n=1}^\infty$ is an infinite sequence of positive numbers to be chosen suitably and $\{x_n\}_{n=1}^\infty$ can be chosen to be any sequence of vectors in H.

Convention: We shall choose the x_n's such that the set $X = \{x_1, x_2, \ldots\}$ is a countable dense set in Ω.

The first lemma gives us estimates for $\|D^j\phi_\rho(x)\|_j$.

Lemma 9.2.

(i) ϕ_ρ *satisfies:* $\forall$ *integers* $j \ge 0$, $\|D^j\phi_\rho(x)\|_j \le \frac{M_j}{\rho^j}$ *for suitable constants* M_j *which are independent of* x, *and* $\rho > 0$.

(ii) The same property holds for $D^j\{1-\phi_\rho(x)\}$, $D^j\phi_\rho(x-c)$ and $D^j\{1-\phi_\rho(x-c)\}$ where $\rho>0$, and c is any fixed point in H, with the same constants as in (i).

Proof of Lemma 9.2. Each derivative $D^j\beta(x)$, of $\beta(\cdot),(j\geq 1)$, is continuous and vanishes outside a compact set viz. $[\frac{1}{2}\leq x\leq 1]$ in $\mathbf{R}^1$, hence is bounded on $\mathbf{R}^1$.

Next, $D(\|x\|^2_{\mathcal{H}}) = F_x$ where F_x is the functional on $\mathcal{H}$ defined by $F_x(h) = 2\langle x,h\rangle$, $x,h\in\mathcal{H}$, with $\|F_x\|_{\mathcal{H}} = 2\|x\|$. Further $D^2(\|x\|^2_{\mathcal{H}}) = G$ where G is the (constant) bilinear functional on $\mathcal{H}\times\mathcal{H}$ defined by $G(h_1,h_2) = 2\langle h_1,h_2\rangle$, $h_1,h_2\in\mathcal{H}$, with $\|G\|_2 = 2$. The further derivatives of $\|x\|^2_{\mathcal{H}}$ are all zero.

Then $D^2\phi(x) = D^j\beta(\|x\|^2_{\mathcal{H}})$ is the sum of a finite number of terms each of which is bounded on $\mathcal{H}$, hence each derivative is bounded on $\mathcal{H}$. Then by the Chain Rule it follows that $\|D\phi_\rho(x)\|_1\leq \frac{M_1}{\rho}$ where $\|D\phi(x)\|\leq M_1$; likewise $\|D^2\phi_\rho(x)\|_2\leq\frac{M_2}{\rho^2}$ where $\|D^2\phi(x)\|_2\leq M_2\forall x\in\mathcal{H}$. This proves the lemma.

We shall make the convention.

Convention. The constants M_j in Lemma 1 will be henceforth chosen as follows. M_0 is chosen to be $=1$, and for $j=1,\ldots,k$, M_j is chosen to be $=\operatorname*{l.u.b}_{x\in\mathcal{H}}\|D^j\phi(x)\|_j$.

Next let $\Phi_n(x) = \phi(x-x)\phi(x-x_2)\ldots\phi(x-x_n)$, and define

$$M_{i,0}(n) = \operatorname*{l.u.b}_{x\in\mathcal{H}}\Phi_n(x),\quad n=1,2,\ldots,\quad\text{and}\quad 0\leq i\leq k\ .$$

Further define

$$M_{j,i}(1) = \binom{j}{i}M_i,\quad 0\leq j\leq k,\ 0\leq i\leq j\ ,$$

where $\binom{s}{t}$ is the binomial coefficient $\frac{s!}{t!(s-t)!}$ for integers $s,t\ni 0\leq t\leq s$. Then define

$$M_{j,i}(2) = \binom{j}{i}\Big[M_{i,i}(1)\cdot M_0 + M_{i,i-1}(1)\cdot M_1+\cdots+M_{i,0}(1)\cdot M_i\Big]\ .$$

Then by Leibnitz's Rule, for $0 \leq j \leq 2, 0 \leq i \leq j$;

$$\underset{x\in\mathcal{H}}{\text{l.u.b.}} \; \|\binom{j}{i} D^i\Phi_2(x)\|_i \leq M_{j,i}(2) \; .$$

Now suppose (finite) $M_{j,i}(p)$ for all integers $p \ni 0 \leq p \leq n-1$ (with $0 \leq j \leq k, 0 \leq i \leq j$) have been defined, satisfying: for integers $0 \leq q \leq p$:

$$\underset{x\in\mathcal{H}}{\text{l.u.b.}} \; \|\binom{j}{i} D^i\Phi_q(x)\|_i \leq M_{j,i}(q) \; .$$

and

$$M_{j,i}(q) \;=\; \binom{j}{i}\Big[M_{i,i}(q-1)\cdot M_0 + M_{i,i-1}(q-1)\cdot M_1 + \cdots + M_{i,0}(q-1)\cdot M_i\Big] \; .$$

Then define

$$M_{j,i}(n) \;=\; \binom{j}{i}\Big[M_{i,i}(n-1)\cdot M_0 + M_{i,i-1}(n-1)\cdot M_1 + \cdots + M_{i,0}(n-1)\cdot M_i\Big] \; .$$

This shows

$$\underset{x\in\mathcal{H}}{\text{l.u.b.}} \; \|\binom{j}{i} D^i\Phi_n(x)\|_i \leq M_{j,i}(n) \quad \text{for} \quad 0 \leq j \leq k, 0 \leq i \leq j \; .$$

Now define: $C(n) = 1 + \max\{M_{j,i}(n) \bigm| 0 \leq j \leq k, 0 \leq i \leq j\}$ for $n = 1, 2, \ldots$. We also note that this preceding inequality concerning $\|D^i\Phi_n(x)\|_i$ still remains true if some or all of the ϕ's are replaced by $(1-\phi)$'s.

III. The proof of Theorem A depends upon suitable local C^∞-approximations in the neighbourhoods of the points x_i, $i = 1, 2, \ldots$, to the given C^k-mapping f. These local C^∞-approximations then have to be put together; however, the customary technique

of partitions of unity could not be used directly in our problem. In our next lemma we obtain such a local approximation in a neighbourhood of each point of Ω; this approximation is even analytic, and is valid in an arbitrary (separable or non-separable) Hilbert space, or even in a Banach space.

Lemma 9.3. *Let $f \in C^k(\Omega, F), x \in \Omega$ and $\eta > 0$. Then $\exists \tilde{f} = \tilde{f}_x : \Omega \to F \ni \tilde{f} \in C^\infty(\Omega, F)$ and $\ni$ in a suitable neighbourhood U of $x, \tilde{f}$ satisfies:*

$$\forall \quad \text{integers} \quad j \in [0,k], \|D^j\tilde{f}(y) - D^j f(y)\|_j < \eta, \quad \forall y \in U \ .$$

Proof of Lemma 9.3. Let $x \in \Omega, \eta > 0$. Then *define* $\tilde{f} = \tilde{f}_x : \Omega \to F$ by

$$\tilde{f}(y) = \tilde{f}_x(y) = f(x) + Df(x) \cdot (y-x) + \frac{D^2 f(x)}{2!} \cdot (y-x)^{(2)} + \cdots + \frac{D^k f(x)}{k!} \cdot (y-x)^{(k)} \ .$$

Here, as is customary, "$(w)^{(j)}$" denotes the j-vector $\overbrace{(w, \ldots, w)}^{j}$. This expression for $\tilde{f}$ is the usual Taylor polynomial of order k of f around the point x. Changing our notation slightly, write

$$\tilde{f}(y) \ = \ A_0 + A_1 \cdot (y-x) + A_2 \cdot (y-x)^{(2)} + \cdots + A_k \cdot (y-x)^{(k)}$$

where $A_0 = f(x) \in F, A_1 = Df(x) \in \mathcal{L}(\mathcal{H}, F)$, $A_2 = \frac{1}{2!}D^2 f(x) \in \mathcal{L}^1_s(\mathcal{H}, F) \ldots$, $A_k = \frac{D^k f(x)}{k!} \in \mathcal{L}^k_s(\mathcal{H}, F)$. Here $\mathcal{L}(\mathcal{H}, F)$ is the space of continuous linear mappings $\mathcal{H} \to F$, and for any integer $j \geq 2$, $\mathcal{L}^j_s(\mathcal{H}, F)$ is the space of j-linear continuous symmetric mappings $\underbrace{\mathcal{H} \times \cdots \times \mathcal{H}}_{j} \to F$.

Here we use the following notation. Suppose $G_m(x_1, \ldots, x_m)$ is an m-linear mapping $\underbrace{V \times \cdots \times V}_{m} \to W$, where V, W are vector spaces. For fixed $x_1, \ldots, x_j \in V$

($1 \le j \le m$), we shall write $G_m(x_1, \ldots, x_j)^{(j)}$ for the mapping $\underbrace{V \times \cdots \times V}_{m-j} \to W$ defined by

$$G_m(x_1, \ldots, x_j)^{(j)}(x_{j+1}, \ldots, x_m) = G_m(x_1, \ldots, x_j, x_{j+1}, \ldots, x_m) .$$

Similarly taking $x_1 = x_2 = \ldots = x_j = x$, we shall write $G_m(x)^{(j)}$ for the mapping $\underbrace{V \times \cdots \times V}_{m-j} \to W$ defined by

$$G_m(x)^{(j)}(x_{j+1}, \ldots, x_m) = G_m\Big(\underbrace{x, \ldots, x_j}_{j}, x_{j+1}, \ldots, x_m\Big) ,$$

and likewise:

$$G_m(x)^{(j)}(x)^{m-j} \stackrel{\text{def}}{=} \Big(\underbrace{x, \ldots, x_j}_{j}, \underbrace{x, \ldots, x}_{m-j}\Big) .$$

A little calculation then shows that (more details will be found in [ ]):

$$\tilde{f}(y) \doteqdot f(x) = A_0;\ D\tilde{f}(y) \doteqdot Df(x) = A_1;\ \ldots,;\ D^k\tilde{f}(y) = D^k f(x) = k!A_k ,$$

provided $y \doteqdot x$. (Here "$\doteqdot$" means "approximately equal to".) Thus $\exists \delta = \delta(x, \eta, f) > 0 \ni$

$$\|y - x\|_{\mathcal{H}} < \delta \Rightarrow \forall \quad \text{integers} \quad j \in [0, k], \|D^j\tilde{f}(y) - D^j f(y)\|_j < \eta .$$

Then it suffices to let $U = U(\delta)$ be the neighbourbood $B_\delta(x)$ of x in Ω. This proves Lemma 9.3.

As a corollary of Lemma 9.3 we obtain:

Corollary 9.4. *The function $\tilde{f}_x$ in the preceding Lemma further satisfies (preserving the notation of the last lemma): $\exists \delta > 0$ satisfying $\forall$ integers $j \in [0, k], \sup_{y \in U} \|D^j\tilde{f}_x(y) - D^j f(y)\|_j < \frac{\eta\delta^{k-j}}{(k-j)!}$.*

Proof of Corollary 9.4. We compare the Taylor expansion (cf. [33], p. 110) of f, or Df, or $D^2 f$, ..., $D^k f$, respectively with $\tilde{f}_x(y)$, or $D\tilde{f}x(y)$, ..., or $D^k\tilde{f}_x(y)$. For any integer $j \in [0,k]$,

$$\begin{aligned}\|D^j\tilde{f}_x(y) - D^j f(y)\|_j &= \Big\|\Big[D^j f(x) + D^{j+1}f(x)\cdot(y-x) + \cdots + \frac{D^k f(x)}{(k-j)!}\cdot(y-x)^{k-j}\Big] \\ &\quad - \Big[D^j f(x) + D^{j+1} f(x)\cdot(y-x) + \cdots \\ &\quad + \int_0^1 \frac{(1-t)^{k-j-1}}{(k-j-1)!} D^k f\big(x + t(y-x)\big)dt\cdot(y-x)^{(k-j)}\Big]\Big\| .\end{aligned}$$

Now suppose $\eta > 0$; then let $\delta > 0 \ni \sup\limits_{y\in U} \|D^k f(x) - D^k f(y)\|_k < \eta$. Then for such y we find that

$$\sup_{y\in U} \|D^j\tilde{f}_x(y) - D^j f(y)\|_j < \frac{\eta\cdot\delta^{k-j}}{(k-j)!}, \qquad j = 0,1,\ldots,k ,$$

This proves the corollary.

We recall that $X = \{x_1, x_2, \ldots\}$ is a countable dense set in Ω. We shall write $\varepsilon_n = \varepsilon(x_n), n = 1,2,\ldots$. We shall now agree on the

Convention. $\varepsilon(\cdot)$ is bounded by 1 on Ω.

For otherwise we can replace $\varepsilon(\cdot)$ by min $\{\varepsilon(\cdot),1\}$.

Lemma 9.5.

(a) *For each $x \in \Omega, \exists$ open ball $B_r(x) \subset \Omega$ satisfying*

$$\sup_{y,y'\in B_r(x)} |\varepsilon(y) - \varepsilon(y')| < \inf_{y\in B_r(x)} \frac{\varepsilon(y)}{2} .$$

(b) *For each $n = 1,2,\ldots,$ and given constant $K_n > 1 \; \exists$ open ball $B_{\rho_n}(x_n) \subset \Omega \ni$ (a) holds in $B_{\rho_n}(x_n)$ as also the following: $\exists \tilde{f}_n \in C^\infty(\Omega, F)$ satisfying:*

$$\forall \text{ integers } \; j \in [0,k], \sup_{x\in B_{\rho_n}(x_n)} \|D^j\tilde{f}_n(x) - D^j f(x)\|_j < \frac{\varepsilon_n\cdot\rho_n^{k-j}}{K_n\cdot 2^{n+3}} .$$

Proof of Lemma 9.5.

(a) Follows because $\varepsilon(\cdot)$ is positive and continuous.

(b) For each $n = 1, 2, \ldots$, $\exists$ open $B_{r_n}(x_n) \ni$ (a) holds in $B_{r_n}(x_n)$. Also by Cor. to Lemma 2, $\exists$ open ball $B_{r'_n}(x_n)$ and $\exists$ function $\tilde{f}_n \in C^\infty(\Omega, F)$ satisfying:

$$\forall \quad \text{integers} \quad j \in [0, k], \quad \sup_{x \in B_{r'_n}(x_n)} \|D^j \tilde{f}_n(x) - D^j f(x)\|_j < \frac{\varepsilon_n \cdot r'^{k-j}_n}{K_n \cdot 2^{n+3}} \,.$$

Now let $\rho_n = \min(r_n, r'_n)$. Then

$$\forall \quad \text{integers} \quad j \in [0, k], \quad \sup_{x \in B_{\rho_n}(x_n)} \|D^j \tilde{f}_n(x) - D^j f(x)\|_j < \frac{\varepsilon_n \cdot \rho^{k-j}_n}{K_n \cdot 2^{n+3}} \,.$$

This proves Lemma 9.5.

From now on, we shall make the following

Convention. The constants K_n in Lemma 3(b) will be chosen to be $= 2(k+1)C(n)$, for $n = 1, 2, \ldots$. Further the ρ_n's shall be $\ni$ both (a) and (b) of Lemma 3 are satisfied in $B_{\rho_n}(x_n)$ and further $\ni \rho_n \geq \rho_{n+1}$ for $n = 1, 2, \ldots$. We shall let $\rho_0 = 1$.

We shall also recall that the function $\phi_{\rho_n}(x - x_n) \in C^\infty\colon H \to [0, 1]$, $\phi_{\rho_n} \equiv 1$ on $B_{\rho_n/\sqrt{2}}(x_n)$ and $\phi_{\rho_n} \equiv 0$ outside $B_{\rho_n(x_n)}$, for each $n = 1, 2, \ldots$.

We now define $\Omega' = \bigcup_{n=1}^{\infty} B_{\rho_n/\sqrt{2}}(x_n)$, and $\Omega'' = \bigcup_{n=1}^{\infty} B_{\rho_n}(x_n)$. The following observation will be needed in the sequel; we shall omit the simple justification.

Lemma 9.6. *If $B_{\rho_m}(x_m) \cap B_{\rho_n}(x_n) \neq \emptyset$, then $\varepsilon_m < 4\varepsilon_n$.*

In the next two lemmas we shall establish convenient bounds for expressions of the type $\operatorname{l.u.b.}_{x \in \mathcal{H}} \|D^j\{\phi_{\rho_1}(x - x_1) \ldots \phi_{\rho_n}(x - x_n)\}\|_j$, or $\|D^j\{\phi_{\rho_1}(x - x_1) \ldots \phi_{\rho_n}(x - x_n)(\tilde{f}_n(x) - f(x))\}\|_j$.

Lemma 9.7. *$\forall$ integers $i, j \ni 0 \le i \le j$, $\underset{x\in\mathcal{H}}{\text{l.u.b.}}\ \|\binom{j}{i}D^i\{\phi_{\rho_1}(x-x_1)\ldots\phi_{\rho_n}(x-x_n)\}\|_i \le \frac{M_{j,i}(n)}{\rho_n^i} \le \frac{C(n)}{\rho_n^i}$. The inequality is still true if some or all of the ϕ's are replaced by $(1-\phi)$'s.*

Proof. We shall use induction on n. For $n = 1$ the statement is true, since (for any $p = 1, 2, \ldots$) we know that for any integer $i \ge 0$, $\underset{x\in\mathcal{H}}{\text{l.u.b.}}\ \|D^i\phi_{\rho_p}(x-x_p)\|_i \le \frac{M_i}{\rho_p^i}$, and $M_{j,i}(1) = \binom{j}{i}M_i$.

Suppose the lemma has been established for some integer $n \ge 1$. Write $\Phi^{\rho_1\cdots\rho_m}(x) = \phi_{\rho_1}(x-x_1)\ldots\phi_{\rho_m}(x-x_m)$ for any m. Then $\Phi^{\rho_1\cdots\rho_{n+1}}(x) = \Phi^{\rho_1\cdots\rho_n}(x)\cdot\phi_{\rho_{n+1}}(x-x_{n+1})$. Using Leibnitz's Rule (cf. II(b)), and the property that $\rho_n \ge \rho_{n+1}$, we find

$$\begin{aligned}\|\binom{j}{i}D^i\Phi^{\rho_1\cdots\rho_{n+1}}(x)\|_i &\le \binom{j}{i}\Bigg[\frac{M_{i,i}(n)}{\rho_n^i}M_0 + \frac{M_{i,i-1}(n)}{\rho_n^{i-1}}\cdot\frac{M_1}{\rho_{n+1}} \\ &\qquad + \cdots + \frac{M_{i,0}(n)}{\rho_n^0}\cdot\frac{M_i}{\rho_{n+1}^i}\Bigg] \\ &\le \frac{\binom{j}{i}}{\rho_{n+1}^i}\Big[M_{i,i}(n)\cdot M_0 + M_{i,i-1}(n)\cdot M_1 + \cdots + M_{i,i}(n)\cdot M_i\Big] \\ &= \frac{M_{j,i}(n+1)}{\rho_{n+1}^i}\end{aligned}$$

This proves the lemma.

Corollary 9.8. *A similar result holds for $\|D^i\{\phi_{\rho_{p+1}}(x-x_{p+1})\ldots\phi_{\rho_{p+m}}(x-x_{p+m})\}\|_i$, viz.*

$$\underset{x\in\mathcal{H}}{\text{l.u.b.}}\ \|\binom{j}{i}D^i\{\phi_{\rho_{p+1}}(x-x_{p+1})\ldots\phi_{\rho_{p+m}}(x-x_{p+m})\}\|_i \le \frac{M_{j,i}(m)}{\rho_{p+m}^i}\,.$$

A similar estimate is valid if some or all of ϕ's are replaced by $(1-\phi)$'s.

Lemma 9.9. *Let $x \in \Omega$. Then for each integer $j \in [0, k]$,*

$$\|D^j\Big\{\phi_{\rho_1}(x-x_1)\ldots\phi_{\rho_n}(x-x_n)\big(\tilde{f}_n(x)-f(x)\big)\Big\}\|_j \le \frac{\varepsilon_n\cdot\rho_n^{k-j}}{2^{n+4}}\,.$$

The inequality remains valid if some or all of the ϕ's except ϕ_{ρ_n} are replaced by $(1-\phi)$'s.

Proof of Lemma 9.9. Using the preceding notation, and Leibnitz's Rule, we find, for $x \in B_{\rho_n}(x_n)$:

$$\begin{aligned}
&\|D^j\Big\{\phi_{\rho_1}(x-x_1)\dots\phi_{\rho_n}(x-x_n)\big(\tilde{f}_n(x)-f(x)\big)\Big\}\|_j \\
&= \|D^j\{\Phi^{\rho_1\dots\rho_n}(x)\cdot\big(\tilde{f}_n(x)-f(x)\big)\}\|_j \\
&\quad \|D^j\{\Phi^{\rho_1\dots\rho_n}(x)\}\cdot\big(\tilde{f}_n(x)-f(x)\big)+\binom{j}{1}\|D^{j-1}\{\Phi^{\rho_1\dots\rho_n}(x)\}\cdot \\
&\qquad\qquad D\big(\tilde{f}_n(x)-f(x)\big)+\dots+\Phi^{\rho_1\dots\rho_n}(x)\cdot D^j\big(\tilde{f}_n(x)-f(x)\big)\|_j \\
&\le \frac{M_{j,j}(n)}{\rho_n^j}\,\frac{\varepsilon_n\rho_n^k}{2^{n+4}(k+1)C(n)}+\dots+\frac{M_{j,0}(n)}{\rho_n^0}\cdot\frac{\varepsilon_n\rho_n^{k-j}}{2^{n+4}\cdot(k+1)C(n)} \\
&\le \frac{(k+1)\cdot\varepsilon_n\rho_n^{k-j}}{2^{n+4}\cdot(k+1)} \\
&= \frac{\varepsilon_n\cdot\rho_n^{k-j}}{2^{n+4}}\,.
\end{aligned}$$

On the other hand, for $x \in \Omega - B_{\rho_n}(x_n)$, the l.h.s. is simply zero, hence the inequality is trivially true. This proves the lemma.

Corollary 9.10. *If $x \in \Omega$, then for any integer $j \in [0,k]$,*

$$\|D^j\Big\{\phi_{\rho_{p+1}}(x-x_{p+1})\dots\phi_{\rho_{p+m}}(x-x_{p+m})\big(\tilde{f}_{p+m}(x)-f(x)\big)\Big\}\|_j \le \frac{\varepsilon_{p+m}\rho_{p+m}^{k-j}}{2^{p+m+4}}\,.$$

The inequality remains valid if some or all of the ϕ's except $\phi_{\rho_{p+m}}$, are replaced by $(1-\phi)$'s.

IV. We now define the functions $g_n : \Omega \to F$, $n = 1, 2, \dots,$ by

$$\begin{aligned}
g_n(x) &= f(x)+\phi_{\rho_1}(x-x_1)\big(\tilde{f}_1(x)-f(x)\big) \\
&\quad +\phi_{\rho_2}(x-x_2)\big(1-\phi_{\rho_1}(x-x_1)\big)\big(\tilde{f}_2(x)-f(x)\big)+\cdots \\
&\quad +\phi_{\rho_n}(x-x_n)\big(1-\phi_{\rho_{n-1}}(x-x_{n-1})\big)\dots\big(1-\phi_{\rho_1}(x-x_1)\big)\big(\tilde{f}_n(x)-f(x)\big)\,.
\end{aligned}$$

This sequence $\{g_n\}$ then has the following properties.

Lemma 9.11. *For each $x \in \Omega'$, the sequence $\{g_n(x)\}$ is constant after a certain stage.*

Proof. If $1 \le m \le n$, then for $x \in B_{\rho_m/\sqrt{2}}(x_m)$,

$$\begin{aligned} g_n(x) &= f(x) + \phi_{\rho_1}(x - x_1)\big(\tilde{f}_1(x) - f(x)\big) + \cdots \\ &\quad + \phi_{\rho_m}(x - x_m)\big(1 - \phi_{\rho_{m-1}}(x - x_{m-1})\big)\ldots\big(1 - \phi_{\rho_1}(x - x_1)\big)\big(\tilde{f}_m(x) - f(x)\big) \\ &= g_m(x) \ . \end{aligned}$$

Lemma 9.12. *On each $B_{\rho_n/\sqrt{2}}(x_n)$, g_n is C^∞.*

Proof. We see that each g_n can be written as

$$\begin{aligned} g_n &= \Big[1 - \phi_{\rho_1} - \phi_{\rho_2}(1 - \phi_{\rho_1}) - \ldots \phi_{\rho_n}(1 - \phi_{\rho_{n-1}})\Big] f = \phi_{\rho_1}\tilde{f}_1 \\ &\quad + \phi_{\rho_2}(1 - \phi_{\rho_1})\tilde{f}_2 + \cdots + \phi_{\rho_n}(1 - \phi_{\rho_{n-1}})\cdots(1 - \phi_{\rho_1})\tilde{f}_n \\ &= (1 - \phi_{\rho_1})(1 - \phi_{\rho_2})\cdots(1 - \phi_{\rho_n})f + \quad \text{(finite number of} \quad C^\infty \text{ terms).} \end{aligned}$$

It follows that on $B_{\rho_n/\sqrt{2}}(x_n)$, g_n is C^∞. Thus the Lemma is proved.

Now define the sequence $\{h^n_p\}, 0 \le p \le n$, by: $h^n_n = 0, h^n_{n-1} = \phi_{\rho_n} \cdot (\tilde{f}_n - f)$; and for $1 \le p \le n-1$, $h^n_{p-1} = \phi_{\rho_p}(\tilde{f}_p - f) + (1 - \phi_{\rho_p})h^n_p$. Then by iteration we find, for $1 \le p \le n$:

$$\begin{aligned} h^n_{p-1} &= \phi_{\rho_p}(\tilde{f}_p - f) + (1 - \phi_{\rho_p})h^n_p = \phi_{\rho_p}(\tilde{f}_p - f) + (1 - \phi_{\rho_p})\phi_{\rho_{p+1}}(\tilde{f}_{p+1} - f) \\ &\quad + (1 - \phi_{\rho_{p+1}})(1 - \phi_{\rho_p})h^n_{p+1} = \cdots = \\ &= \phi_{\rho_p}(\tilde{f}_p - f) + \phi_{\rho_{p+1}}(1 - \phi_{\rho_p})(\tilde{f}_{p+1} - f) + \cdots \\ &\quad + \phi_{\rho_n}(1 - \phi_{\rho_{n-1}})\cdots(1 - \phi_{\rho_1})(\tilde{f}_n - f) \ . \end{aligned}$$

In particular,

$$\begin{aligned} h^n_0 &= \phi_{\rho_1}(\tilde{f}_1 - f) + \phi_{\rho_2}(1 - \phi_{\rho_1})(\tilde{f}_2 - f) + \cdots \\ &\quad + \phi_{\rho_n}(1 - \phi_{\rho_{n-1}})\cdots(1 - \phi_{\rho_1})(\tilde{f}_n - f) \ , \end{aligned}$$

and hence $g_n = f + h^n_0$.

Lemma 9.13. *For any integer $j \in [0,k]$,* $\sup_{x \in B_{\rho_n/\sqrt{2}}(x_n)} \|D^j g_n(x) - D^j f(x)\| j < \frac{\varepsilon_n}{2}$.

Proof of Lemma 9.13. We shall show that for any integer $j \in [0,k]$, and any integer $p \in [1,n]$,

$$\sup_{x \in B_{\rho_n/\sqrt{2}}(x_n)} \|D^j h^n_{p-1}(x)\| < \frac{\varepsilon_n \rho_{p-1}^{k-j}}{2^{p+1}} .$$

Let $1 \leq p \leq n-1$; and suppose $\Big(B_{\rho_p}(x_p) \cap B_{\rho_{p+1}}(x_{p+1}) \cap \ldots \cap B_{\rho_{n-1}}(x_{n-1})\Big) \cap B_{\rho_n/\sqrt{2}}(x_n)$ $\neq \emptyset$. Then for $x \in \Big(B_{\rho_p} \cap \ldots \cap B_{\rho_{n-1}}(x_{n-1})\Big) \cap B_{\rho_n/\sqrt{2}}(x_n)$, by using Lemmas 9.7 and 9.9 we find that

$$\begin{aligned} \|D^j h^n_{p-1}(x)\|_j &\leq \frac{\varepsilon_p \rho_p^{k-j}}{2^{p+4}} + \frac{\varepsilon_{p+1} \rho_{p+1}^{k-j}}{2^{p+5}} + \cdots + \frac{\varepsilon_n \rho_n^{k-j}}{2^{n+4}} \leq \frac{\rho_{p-1}^{k-j}}{2^{p+4}} \cdot 8\varepsilon_n \\ &= \frac{\varepsilon_n \rho_{p-1}^{k-j}}{2^{p+1}} . \end{aligned}$$

On the other hand if $B_{\rho_m}(x_m) \cap B_{\rho_n/\sqrt{2}}(x_n) = \emptyset$ for some $m \leq n-1$, then for $x \in B_{\rho_n/\sqrt{2}}(x_n)$ and $\ni x \notin B_{\rho_m}(x_m)$ for $m \leq n-1$, the corresponding term in the preceding finite sum is to be replaced by 0, and hence the estimate is again valid. Hence for $x \in B_{\rho_n/\sqrt{2}}(x_n)$

$$\|D^j h^n_0(x)\|_j = \|D^j g_n(x) - D^j f(x)\|_j < \frac{\varepsilon_n \rho_0^{k-j}}{4} = \frac{\varepsilon_n}{4} ,$$

hence $\sup_{x \in B_{\rho_n/\sqrt{2}}(x_n)} \|D^j g_n(x) - D^j f(x)\|_j < \frac{\varepsilon_n}{2}$. This completes the proof of Lemma 9.

Now we define $\bar{g} = \lim_{n->\infty} g_n$ on Ω'. For each $x \in \Omega'$ set $n_x = \inf\Big[m \;\Big|\; x \in B_{\rho_m/\sqrt{2}}(x_m)\Big]$. Then $\exists$ neighbourhood $V(x)$ of $x \ni V(x) \subset B_{\rho_{n_*}/\sqrt{2}}(x_{n_*})$, and $\ni$ the integer $n.$ is constant in $V(x)$ or changes to a smaller value in which case we use the result: if $x \in B_{\rho_m/\sqrt{2}}(x_m)$ then for any integer $n \geq m, g_n(x) = g_m(x)$. Therefore we conclude that $\bar{g}$ ic C^∞ on Ω' and for each $x \in \Omega'$:

$$\forall \quad \text{integers} \quad j \in [0,k], \quad \sup_{y \in B_{\rho_n/\sqrt{2}}(x_{n_*}) \cap V(x)} \|D^j \bar{g}(y) - D^j f(y)\|_j < \frac{\varepsilon(x_{n_*})}{2}$$

$$< \varepsilon(y) \quad \forall \quad y \in B_{\rho_{n_*}/\sqrt{2}}(x_{n_*}) \cap V(x) \ .$$

In the next lemma we shall show that $\lim_{n\to\infty} g_n(x)$ exists not only on Ω' but exists, in fact uniformly, for all $x \in \Omega$.

Lemma 9.14. *For each integer $j \in [0,k]$, the sequence $\{D^j g_n(x)\}$ is uniformly convergent $\forall x \in \Omega$. Hence the function $\bar{g}$ defined by: $\bar{g}(x) = \lim_{n\to\infty} g_n(x)$ at each point $x \in \Omega$ is C^k-smooth in Ω and has the approximation property:*

$$\forall \quad \text{integers} \quad j \in [0,k], \|D^j\bar{g}(x) - D^j f(x)\|_j < \varepsilon(x) \qquad \forall x \in \Omega \ .$$

Proof of Lemma 9.14. Let N, N' be integers $\ni N' > N > 0$. Then for any integer $j \in [0,k]$, and $x \in \Omega$, by Lemma 9.9:

$$\begin{aligned} \|D^j g_{N'}(x) - D^j g_N(x)\|_j &\le \frac{\varepsilon_{N+1}\rho_{N+1}^{k-j}}{2^{N+5}} + \cdots + \frac{\varepsilon_{N'}\rho_{N'}^{k-j}}{2^{N'+4}} \\ &< \frac{1}{2^{N+5}} + \cdots + \frac{1}{2^{N'+4}} \to 0 \quad \text{as} \quad N, N' \to \infty \ . \end{aligned}$$

Hence for such j the sequence $\{D^j g_n(x)\}$ is uniformly convergent $\forall x \in \Omega$. For $j = 0$ this means $\{g_n(\cdot)\}$ is uniformly convergent in Ω. Hence $\bar{g}(x) = \lim_{n\to\infty} g_n(x)$ exists not merely at points in Ω' but at all points in Ω. Furthermore by Theorem 12 in [33], p. 117, we are able to assert that for each integer $j \in [1,k]$, $\lim_{n\to\infty} D^j g_n$ must be $= D^j\bar{g}$, and further since $\{D^j g_n\}$ is uniformly convergent and each $D^j g_n$ is continuous in Ω, therefore $D^j\bar{g}$ is continuous in Ω. Thus $\bar{g}$ is C^k-smooth in Ω.

Next for such integers j, and each point $x_m \in X$,

$$\begin{aligned} \|D^j\bar{g}(x_m) - D^j f(x_m)\|_j &= \|D^j g_{n_{*m}}(x_m) - D^j f(x_m)\|_j \\ &= \|D^j g_m(x_m) - D^j f(x_m)\|_j \end{aligned}$$

because $n_{x_m} \leq m$; and hence

$$\|D^j \bar{g}(x_m) - D^j f(x_m)\|_j < \frac{\varepsilon_m}{2} .$$

Let $y \in (\partial\Omega') \cap \Omega$ if $(\partial\Omega') \cap \Omega \neq \emptyset$. Then $\exists$ subsequence $\{x_{n_m}\} \ni x_{n_m} \to y$ and because $D^j\bar{g}(\cdot), D^j f(\cdot), \varepsilon(\cdot)$ are all continuous on Ω, we find:

$$\|D^j \bar{g}(y) - D^j f(y)\|_j \leq \frac{\varepsilon_m}{2} < \varepsilon(y) .$$

Hence $\bar{g}$ has the property:

$$\forall \quad \text{integers} \quad j \in [0, k], \|D^j \bar{g}(x) - D^j f(x)\|_j < \varepsilon(x) \forall x \in \Omega .$$

This completes the proof of Lemma 9.14 and hence Theorem A is proved.

V. We shall now turn to the proof of Theoremm B stated at the beginning of this section. We shall need the following two lemmas. Although these two lemmas are stated for the particular functions $g(\cdot)$ and $\sigma(\cdot)$, they are clearly true more generally.

Lemma 9.15. *The function $\gamma(\cdot)$ defined for $x \in \Omega$ by:*

$$\gamma(x) = \sup\Big[0 < \delta < 1 \;\Big|\; \underset{\delta}{B}(x) \subset \Omega, \sup_{y,y' \in \underset{\delta}{B}(x)} \|\bar{g}(y) - \bar{g}(y')\|_F < \inf_{y \in B_\delta(x)} \frac{\sigma(y)}{2^{k+1}}\Big] ,$$

is continuous on Ω.

Proof of Lemma 9.15. Let $x \in \Omega$, and let $\{x_n\}_{n=1}^{\infty}$ be a sequence of points in Ω converging to x. We first show that $\bar{\gamma} = \limsup_{n\to\infty} \gamma(x_n) \leq \gamma(x)$; by similar arguments it will follow that $\underline{\gamma} = \liminf_{n\to\infty} \gamma(x_n) \geq \gamma(x)$.

Taking a subsequence of $\{x_n\}$ if necessary we shall suppose $\lim_{n\to\infty} \gamma(x_n) = v$. Suppose $v > \gamma(x)$. Then a simple argument shows that for n sufficiently large, $B_{\gamma(x_n)}(x_n)$ would properly contain $\overline{B_{\gamma(x)}(x)}$. However this would clearly violate the supremum property of $\gamma(x)$. We therefore conclude that v must be $\leq \gamma(x)$. Hence $\bar{\gamma}$ must be $\leq \gamma(x)$.

Similarly we conclude that $\underline{\gamma} \geq \gamma(x)$, because otherwise the supremum property of $\gamma(x_n)$ would be violated for many n's. This completes the proof of the lemma.

The proof of the next lemma is very similar to that of Lemma 9.15, and hence will be omitted.

Lemma 9.16. *The function $\lambda(x)$ defined for $x \in \Omega$ by*

$$\lambda(x) = \sup\Bigg[0 < \delta < 1 \;\Bigg|\; \underset{\delta}{B}(x) \subset \Omega, \quad \forall \quad \text{integers} \quad j \in [0,k],$$
$$\sup_{y,y' \in \underset{\delta}{B}(x)} \|D^j \bar{g}(y) - D^j \bar{g}(y')\|_j < \inf_{y \in \underset{\delta}{B}(x)} \frac{\sigma(y)}{2^{k+1}}\Bigg],$$

is continuous on Ω.

We now have to define two functions $\mu(x)$ and $N(x)$. For $x \in \Omega$, let

$$\mu(x) = \frac{1}{2N(x)} \quad \min\{\lambda(x), \sigma(x)\},$$

where

$$N(x) = 1 + \max_{0 \leq j \leq k} \left(\|D^j \bar{g}(x)\|_j\right).$$

The preceding lemmas and the proof of Theorem A show that these functions are positive and continuous on Ω.

We shall now introduce suitable C^∞ mappings $u_\alpha(\cdot) : \Omega \to H$ which map suitable subsets of Ω into Ω'. The composite mappings $\bar{g}(u_\alpha(\cdot))$ will enable us to define a

suitable C^∞ mapping $g(\cdot) : \Omega \to F$. Each of the mappings $u_\alpha(\cdot)$ will be defined with the help of an element of the dense countable set X; further it will map a neighbourhood $B_r(x)$ of some point $x \in \Omega$ into Ω' and will have the properties: for $x' \in B_r(x)$, $\|u_\alpha(x') - x'\|_{\mathcal{H}} < \frac{\mu(x')}{2}$, $\|D(u_\alpha(x') - x')\|_1 < \frac{\mu(x')}{2}$, and $D^j u_\alpha(x') = 0$ for $j = 2, 3, \ldots$.

Let $x \in \Omega$. Let $\delta' \ni$

(a) $0 < \delta' < 1$

(b) $B_{\delta'}(x) \subset \Omega$, and

(c) $x' \in \underset{\delta'}{B}(x) \Rightarrow |\mu(x') - \mu(x)| < \frac{\mu(x)}{4}$;

and define

$$\text{(i) } \nu = \frac{3}{16}\mu(x); \qquad \text{(ii) } t = \frac{1}{m}$$

where $m > 1$ is the unique integer $\ni$

$$\frac{1}{m} < \frac{1}{2}\frac{\log(\frac{1}{1-\nu})}{\log 2} \leq \frac{1}{m-1} ,$$

the logarithms being natural logarithms. Let x_{n_p} be the first element of X, different from x, which lies in $B_{\frac{\mu(x)}{2}}(x)$. To each $x' \in \underset{\delta'}{B}(x)$ we assign x_{n_p} as the required element of X which we seek. Next let

$$r = \frac{1}{2^{t+1}}\|x - x_{n_p}\|_{\mathcal{H}}$$

and

$$r' = (1 - \frac{1}{2^t})r = (1 - \frac{1}{2^t})\frac{1}{2^{t+1}} \cdot \|x - x_{n_p}\|_{\mathcal{H}} .$$

Then let $\delta'' \leq \min\left[\delta'; \frac{1}{2}(1 - \frac{1}{2^t})(1 - \frac{1}{2^{t+1}}) \cdot \|x - x_{n_p}\|_{\mathcal{H}}\right]$, such that if we define, for each $x' \in B_{\delta''}(x)$ the point z' by:

$$z' = z'_{x'} = (1 - \frac{1}{2^t})x_{n_p} + \frac{1}{2^t}x' ,$$

then $z'_{x'} \in B_{\frac{r'}{16}}(z'_x)$. Now if the positive number r'' is defined by:

$$r'' = \frac{r'}{16},$$

then for every such z', $B_{r''}(z') \subset B_{\frac{r'}{8}}(z'_x)$. This choice of δ'' also ensures that $B_{\delta''}(x) \cap B_{r'}(z'_x) = \emptyset$.

Now let $x_{n_{p'}}$ be the first element of X which lies in $B_{r'}(z'_x) - \overline{B_{\frac{r'}{8}}(z'_x)}$. For each $x' \in B_{\delta''}(x)$, determine the point $z = z_{x'} \in B_r(x_{n_p})$ such that the two triangles, one with vertices $x', z'_{x'}, x_{n_{p'}}$, and the other with vertices x', x_{n_p}, z, are similar. (The relevant geometrical property in a two-dimensional plane can be deduced as a proposition in analysis, see [19]. If dimension of $\mathcal{H}$ equals 1, our arguments need only slight modification.)

Because of the one-one bi-continuous relation between the pairs $(x', z'_{x'})$ with $x' \in B_{\delta''}(x)$ ont the one hand, and the points $z = z_{x'}$ determined by the above process on the other hand, we find that the set Z of all such points z is an open set.

Now let $\eta' = \min\left[\frac{1}{2}\{\|z'_x - x_{n_{p'}}\|_{\mathcal{H}} - \frac{r'}{8}; \text{dist }(x_{n_{p'}}, \partial\Omega'); \frac{1}{2}\{r' - \|z'_x - x_{n_{p'}}\|_{\mathcal{H}}\right]$, so that $B_{\eta'}(x_{n_{p'}}) \subset \Omega'$. Let $\zeta' \ni \frac{\eta'}{\zeta'} = (1 - \frac{1}{2^t})$. Because Z is an open set, therefore $\exists$ open ball $B_\zeta(z_x) \subset Z$, with $\zeta \leq \min\left[\zeta'; \frac{1}{2}\{\|x_{n_p} - z_x\|_H\right]; \text{dist }(z_x, \partial Z)\Big]$. Let x_{n_q} be the first element of X, different from z_x, lying in $B_{\zeta/4}(z_x) \cap Z$.

Finally let $\eta' = (1 - \frac{1}{2^t})\zeta$, and $2\delta = \min\left[\frac{1}{2^t}\eta; \delta''\right]$, where $\eta = \frac{1}{2}\{\eta^1 - \|u(x) - x_{n_{p'}}\|_{\mathcal{H}}\}$, and $u(x) = (1 - \frac{1}{2^t})x_{n_q} + \frac{1}{2^t}x$. Now define, for each $x' \in B_{2\delta}(x)$:

$$u(x') = (1 - \frac{1}{2^t})x_{n_q} + \frac{1}{2^t}x'$$

This mapping $u(\cdot)$ which maps the neighbourhood $B_{2\delta}(x)$ in Ω into a neighbourhood of $u(x)$ in Ω', is, however, defined everywhere (x_{n_q} being fixed), and $\forall x'$: $\|u(x') -$

$x'\|_{\mathcal{H}} = \frac{1}{2^t}\|x_{n_q} - x'\|_{\mathcal{H}}$; $\|D(u(x') - x')\|_1 = 1 - \frac{1}{2^t}$; and $D^j u(x') = 0$ for $j \geq 2$. Furthermore for $x' \in B_{2\delta}(x)$:

$$\begin{aligned} u(x') \in \Omega; \|u(x') - x'\|_{\mathcal{H}} = (1 - \frac{1}{2^t})\|x_{n_q} - x'\|_{\mathcal{H}} < \frac{\mu(x')}{2} \\ \|D(u(x') - x')\|_{\mathcal{H}} = 1 - \frac{1}{2^t} < \frac{\mu(x')}{2}, \text{ and} \\ D^j u(x') = 0 \quad \text{for } j \geq 2 . \end{aligned} \tag{1}$$

It will be convenient to change our notation slightly. We find from the preceding that for each $z \in \Omega$ $\exists$ point $q_z \in X$ and $\exists C^\infty$ function $u(\cdot)$: $\mathcal{H} \to \mathcal{H}$ determined by the above process with the help of q_z, satisfying: for some $t > 0$, and $\forall x$:

$$\|u(x) - x\|_{\mathcal{H}} = (1 - \frac{1}{2^t})\|x - q_z\|_{\mathcal{H}}; \qquad \|D(u(x) - x)\|_1 = 1 - \frac{1}{2^t};$$

$$\|D^j u(x) = 0 \text{ for } j = 2, 3, \dots .$$

Furthermore $\exists \delta > 0 \ni \forall x \in B_{2\delta}(z)$:

$$\begin{aligned} u(x) \in \Omega'; \quad \|u(x) - x\|_{\mathcal{H}} &= (1 - \frac{1}{2^t})\|x - q_z\|_{\mathcal{H}} < \frac{\mu(x)}{2}; \\ \|D(u(x) - x)\|_1 = 1 - \frac{1}{2^t} &< \frac{\mu(x)}{2}; \quad \text{and} \\ D^j u(x) = 0 \quad &\text{for} \quad j \geq 2 . \end{aligned} \tag{2}$$

The separability of $\mathcal{H}$ implies:

(i) $\exists$ countable collection $\{z_1, z_2, \dots\}$ of points in Ω;

(ii) for each $z_n \exists$ a point $q_{z_n} \in X$, a number $\delta_n > 0$, with $\Omega = \bigcup_{n=1}^{\infty} B_{\delta_n/\sqrt{2}}(z_n)$, and a C^∞ function $u_n(\cdot)$: $\mathcal{H} \to \mathcal{H}$ defined with the help of q_{z_n} with properties similar to those in (2), viz. $\forall x \in B_{2\delta_n}(z_n)$:

$$\left.\begin{aligned} &u_n(x) \in \Omega'; \\ &\|u_n(x) - x\|_{\mathcal{H}} = (1 - \frac{1}{2^{t_n}})\|x - q_{z_n}\|_{\mathcal{H}} < \frac{\mu(x)}{2} \quad \text{for some } t_n > 0; \\ &\|D(u_n(x) - x)\|_{\mathcal{H}} = (1 - \frac{1}{2^{t_n}}) < \frac{\mu(x)}{2}; \\ &\text{and} \quad D^j u_n(x) = 0 \quad \text{for} \quad j \geq 2 . \end{aligned}\right\}$$

For each $n = 1, 2, \dots,$ the composite function $\bar{g}(u_n(\cdot))$ is C^∞ in $B_{2\delta_n}(z_n)$ because $\bar{g}(\cdot)$ has been shown to be C^∞ on Ω' and the C^∞ function $u_n(\cdot) : \mathcal{H} \to \mathcal{H}$ maps $B_{2\delta_n}(z_n)$ into Ω'. Further we note that $D^j\bar{g}(\cdot)$ as also $D^j\bar{g}(u_n(\cdot))$, for $j = 0, 1, \dots, k$, is bounded on $B_{2\delta_n}(z_n)$ for $n = 1, 2, \dots$.

Now for each $n = 1, 2, \dots,$ let $\phi_n(\cdot) = \phi_{\delta_n}(\cdot - z_n)$ be the C^∞-function mapping $\mathcal{H} \to [0,1] \ni \phi_n \equiv 1$ on $\overline{B_{\delta_n/\sqrt{2}}}(z_n)$ and $\phi_n \equiv 0$ outside $B_{2\delta_n}(z_n)$ (cf. subsection II(c) above).

VI. We now want estimates for $\|D^j\{\bar{g}(u_n(x)) - \bar{g}(x)\}\|_j$, $j = 0, 1, \dots, k$, $x \in B_{2\delta_n}(z_n)$, $n = 1, 2, \dots$. For this computation we shall use the following known formula which is verified by induction (cf. [12] p. 222): for $x \in B_{2\delta_n}(z_n)$, and $j \geq 1$,

$$\begin{aligned} \|D^j\{\bar{g}(u_n(x)) &= D^j\bar{g}(u_n) \circ \left((Du_n(x))^{(j)}\right) \\ &+ \sum_{\substack{\alpha \in S_j \\ \Sigma\alpha \leq j-1}} A_\alpha D^{\Sigma\alpha}\bar{g}(u_n) \circ \left((Du_n(x))^{(\alpha_1)}, (D^2u_n(x))^{(\alpha_2)}, (D^ju_n(x))^{(\alpha_j)},\right) \end{aligned} \tag{3}$$

where

(i) "$(w)^{(j)}$" means "$\overbrace{(w, \dots, w)}^{j}$";

(ii) $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_j)$, the α_i being non-negative integers satisfying: $\alpha_1 + 2\alpha_2 + \cdots + j\alpha_j = j$,

(iii) $\sum \alpha \stackrel{\text{def}}{=} \alpha_1 + \alpha_2 + \cdots + \alpha_j$;

(iv) for each integer j $(1 \leq j \leq k)$, α further satisfies (in addition to (ii)):

$$\sum \alpha = p_i$$

with $p_i = 1, 2, \dots, j$, i.e. for each j (with $1 \leq j \leq k$) we consider possible solutions

of each of j pairs of simultaneous equations:

$$\begin{pmatrix} \sum\alpha = 1 \\ \sum_{i=1}^{j} i\alpha_i = j \end{pmatrix}; \quad \begin{pmatrix} \sum\alpha = 2 \\ \sum_{i=1}^{j} i\alpha_i = j \end{pmatrix}; \ \ldots \ ; \begin{pmatrix} \sum\alpha = j \\ \sum_{i=1}^{j} i\alpha_i = j \end{pmatrix};$$

(v) A_α are certain positive constants (the numerical values of which do not matter in our calculations);

(vi) "$\alpha \in S_j$" simply means: "the j-tuplet $\alpha = (\alpha_1, \ldots, \alpha_j)$ satisfies the preceding conditions (ii)-(iv)".

In the preceding formula (3) $\Sigma\alpha$ is given the successive values $j, j-1, \ldots, 1$. The first term in the summation on the right side in (3) then corresponds to the unique solution: $\alpha_1 = j$, $\alpha_2 = 0 = \ldots = \alpha_j$ of the simultaneous equations

$$\left.\begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_j &= j, \\ \alpha_1 + 2\alpha_2 + \cdots + j\alpha_j &= j. \end{aligned}\right\} \tag{4}$$

The last term in the summation in (3) is $D\bar{g}(u_n) \circ \big(D^j u_n(x)\big)$, corresponding to the unique solution $\alpha_1 = 0 = \ldots = \alpha_{j-1}, \alpha_j = 1$ of the simultaneous equations

$$\left.\begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_j &= 1, \\ \alpha_1 + 2\alpha_2 + \cdots + j\alpha_j &= j. \end{aligned}\right\} \tag{5}$$

The remaining terms in the summation for $D^j\bar{g}\big(u_n(x)\big)$ correspond to solutions in non negative integers α_i of the pairs of simultaneous equations

$$\left.\begin{aligned} \alpha_1 + \cdots + \alpha_j &= i, \\ \alpha_1 + 2\alpha_2 + \cdots + j\alpha_j &= j. \end{aligned}\right\} \tag{6}$$

with $i = 2, 3, \ldots, j-1$ in turn. In each of these cases, *any* solution α (for a given i with $2 \le i \le j-1$) will contain a non-zero α_i with $\alpha_i > 1$. The corresponding contribution to the summation representing $D^j\bar{g}\big(u_n(x)\big)$ is zero (if $j > 1$).

Then for a given j (with $1 \le j \le k$):

$$\begin{aligned}
\|D^j\bar{g}\big(u_n(x)\big) - D^j\bar{g}(x)\|_j &\le \|D^j\bar{g}(u_n) \circ \Big(\big(Du_n(x)\big)^{(j)}\Big) - D^j\bar{g}(x) \circ \Big(\big(Du_n(x)\big)^{(j)}\Big)\|_j \\
&\quad + \|D^j\bar{g}(x) \circ \Big(\big(Du_n(x)\big)^{(j)} - (I, \dots, I)\Big)\|_j \\
&\quad + \sum_{\substack{\alpha \in S_j \\ \Sigma\alpha \le j-1}} A_\alpha D^{\Sigma\alpha}\bar{g}(u_n) \circ \Big(\big(Du_n(x)\big)^{(\alpha_1)}, \dots, \big(D^j u_n(x)\big)^{(\alpha_j)},\Big)\|_{\Sigma\alpha} \\
&\le 2^j \cdot \frac{\sigma(x)}{2^{k+1}} + N(x) \cdot \big(\mu(x)\big)^j \le \frac{\sigma(x)}{2} + N(x) \cdot \mu(x) \le \frac{\sigma(x)}{2} + N(x) \cdot \frac{\sigma(x)}{2^{N(x)}} \\
&= \frac{\sigma(x)}{2} + \frac{\sigma(x)}{2} = \sigma(x) \ .
\end{aligned}$$

This estimate clearly also holds for $j = 0$.

VII. We now want to put together the composites $\bar{g}\big(u_n(\cdot)\big)$, $n = 1, 2, \dots$, in a convenient manner, to obtain a suitable C^∞ mapping. We proceed as follows.

We define a C^∞ function $h(\cdot) : \Omega \to [0, 1]$ in the following manner. Let $\delta_{(1)} = \delta_1$, and generally for any integer $n \ge 1$, $\delta_{(1,\dots,n)} = \min(\delta_1, \dots, \delta_n)$. Let $x \in \Omega$, and define for $m = 1, 2, \dots$,

$$\begin{aligned}
h_m(x) &= \frac{\delta^k_{(1)}}{2(k+2)C(1)}\phi_1(x) + \frac{\delta^k_{(1,2)}}{2^2(k+2)C(2)}\phi_2(x)\big(1 - \phi_1(x)\big) + \cdots \\
&\quad + \frac{\delta^k_{(1,\dots,m)}}{2^m(k+2)C(m)}\phi_m(x)\big(1 - \phi_{m-1}(x)\big) \cdots \big(1 - \phi_1(x)\big) \ .
\end{aligned}$$

Here the constants $C(j)$ are the ones defined in subsection II(c). For each $m = 1, 2, \dots$, $h_m(\cdot) \in C^\infty(\Omega, \mathbf{R}^1)$. Further the sequence $\big\{h_m(x)\big\}_{m=1}^\infty$ is constant from $m = n$ onward, where for $x \in \Omega$, $n = n_x = \inf\big[m \;\big|\; x \in B_{\delta_m/\sqrt{2}}(z_m)\big]$. Hence $h(x) = \lim_{m \to \infty} h_m(x)$ exists for each $x \in \Omega$, and the function $h(\cdot) \in C^\infty(\Omega, \mathbf{R}^1)$.

We want to note further properties of the function $h(\cdot)$. As noted, if $x \in \Omega$, and

$n = n_x = \inf\left[m \;\middle|\; x \in B_{\delta_m/\sqrt{2}}\right]$ then $h(x) = h_n(x)$. Write $h(x)$ as

$$h(x) = a_1\phi_1(x) + a_2\phi_2(x)\big(1 - \phi_1(x)\big) + \cdots + a_n\phi_n(x)\big(1 - \phi_{n-1}(x)\big)\cdots\big(1 - \phi_1(x)\big)$$

where $a_i = \frac{\delta^k_{(1,\ldots,i)}}{2^i(k+2)C(i)}$. Then $a_m > a_{m+1}$ for any $m = 1, 2, \ldots$. Hence

$$\begin{aligned} h(x) \geq & a_n\Big[\phi_1(x) + \phi_2(x)\big(1 - \phi_1(x)\big) + \cdots + \phi_n(x)\big(1 - \phi_{n-1}(x)\big)\cdots\big(1 - \phi_1(x)\big)\Big] \\ = & a_n\Big[1 - \big(1 - \phi_1(x)\big)\big(1 - \phi_2(x)\big)\cdots\big(1 - \phi_n(x)\big)\Big] = a_n > 0\ . \end{aligned}$$

This means that $\frac{1}{h(x)} \in C^\infty(\Omega, \mathbf{R}^1)$.

With the notation of the preceding paragraphs, for $j = 0, 1, \ldots, k$, and making use of Lemma 9.9 (subsection II(c)), we find

$$\begin{aligned} &\|D^j h(x)\|_j + \binom{j}{1}\|D^{j-1}h(x)\|_{j-1} + \cdots + h(x) \\ &\leq \frac{1}{k+2}\left\{\frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n}\right\} + \frac{1}{k+2}\left\{\frac{1}{2} + \cdots \frac{1}{2^n}\right\} + \cdots + \frac{1}{k+2}\left\{\frac{1}{2} + \cdots \frac{1}{2^n}\right\}\ , \end{aligned}$$

there being no more than $(k+1)$ terms in the sum in the right side. Thus, for $x \in \Omega$, and $j = 0, 1, \ldots, k$,

$$\|D^j h(x)\|_j + \binom{j}{1}\|D^{j-1}h(x)\|_{j-1} + \cdots + \leq \rho\Big[\frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n}\Big]$$

where $\rho = \frac{k+1}{k+2}$, $n = n_x$ being as defined earlier.

Now define $\tilde{g} : \Omega \to F$, by

$$\begin{aligned} \tilde{g}(x) \ &= a_1\phi_1(x)\bar{g}\big(u_1(x)\big) + a_2\phi_2(x)\big(1 - \phi_1(x)\big)\bar{g}\big(u_2(x)\big) + \cdots + \\ &\quad + a_n\phi_n(x)\big(1 - \phi_{n-1}(x)\big)\cdots\big(1 - \phi_1(x)\big)\bar{g}\big(u_n(x)\big)\ , \end{aligned}$$

$n = n_x$ being as defined above, and then define $g(x) = \frac{1}{h(x)}\tilde{g}(x)$. Clearly both $\tilde{g}$ and g

are $C^\infty : \Omega \to F$. Then for $j = 0, 1, \ldots, k$, $x \in \Omega$ and $n = n_x$ as above,

$$\begin{aligned}
\|D^j\Big\{h(x)\big(\bar{g}(x) - g(x)\big)\Big\}\|_j &= \|D^j\{a_1\phi_1(x)\}\Big(\bar{g}(x) - \bar{g}(u_1(x))\Big) \\
&\quad + a_2\phi_2(x)\big(1 - \phi_1(x)\big)\Big(\bar{g}(x) - \bar{g}\big(u_2(x)\big)\Big) \\
&\quad + \cdots + a_n\phi_n(x)\big(1 - \phi_{n-1}(x)\big)\cdots\big(1 - \phi_1(x)\big)\Big(g(x) - \bar{g}\big(u_n(x)\big)\|_j \qquad (7) \\
&\le \frac{1}{k+2}\left[\frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n}\right] \cdot \max_{\substack{0 \le j \le k \\ 1 \le i \le n}} \|D^j\Big\{\bar{g}(x) - \bar{g}\big(u_i(x)\big)\Big\}\|_j \\
&< \rho\left[\frac{1}{2} + \cdots \frac{1}{2^n}\right] \cdot \sigma(x)\ .
\end{aligned}$$

In general, induction yields: if m is any integer ≥ 1, then

$$\|D^j\{h^m(x)\bar{g}(x) - h^m(x)g(x)\}\|_j < \rho^m\left[\frac{1}{2} + \cdots \frac{1}{2^n}\right] \cdot \sigma(x)\ .$$

This suggests that such an estimate might be valid for $m = 0$. We verify that this actually is the case, as follows. For $j = 0, 1, \ldots, k$,

$$\begin{aligned}
\|D^j\{h(x)\big(\bar{g}(x) - g(x)\big)\}\|_j &= \|D^j h(x) \cdot \big(\bar{g}(x) - g(x)\big)\|_j \\
&\quad + \binom{j}{1}D^{j-1}h(x) \cdot D\big(\bar{g}(x) - g(x)\big) + \cdots + h(x)D^j\big(\bar{g}(x) - g(x)\big)\|_j \\
&\le \Big\{\max_{0 \le i \le j} \|D^i\big(\bar{g}(x) - g(x)\big)\|_i\Big\}\Big[\|D^j h(x)\|_j + \binom{j}{1}\|D^{j-1}h(x)\|_{j-1} + \cdots + h(x)\Big], \qquad (8)
\end{aligned}$$

and, on the other hand, the sum in the left side in (8) is

$$\le \sigma(x) \cdot \Big[\|D^j h(x)\|_j + \binom{j}{1}\|D^{j-1}h(x)\|_{j-1} + \cdots h(x)\Big]\ .$$

We now need the following lemma.

Lemma 9.17. *Let $W_1 = \big\{w = (w_j, w_{j-1}, \ldots, w_0)\,\big|\, w_i \in L_s^i(H, \mathbf{R}^1)\big\}$, where j is a fixed integer $\in [0, k]$, with norm $\|w\| = \sum_{i=0}^{j} \|w_i\|_i$. Define T linear: $W_1 \to L_s^j(H, F)$ by*

$$Tw \;=\; w_j \cdot c_0 + w_{j-1} \cdot c_1 + \cdots + w_0 \cdot c_j$$

where $c_i \in L^i_s(H,F)$ are fixed, the products being tensor products. Then $\|T\| = \max\{\|c_i\|_i,\ 0 \le i \le j\}$.

Proof of Lemma 9.17.

$$(i) \qquad \|Tw\| = \|w_j \cdot c_0 + \cdots + w_0 c_j\|_j \le \|w_j\|_j \cdot \|c_0\|_0 + \cdots + \|w_0\|_0 \cdot \|c_j\|_j$$
$$\le \big(\|w_j\|_j + \cdots + \|w_0\|_0\big) \cdot \max_{0\le i\le j} \|c_i\|_i = \|w\| \max_{0\le i\le j} \|c_i\|_i \, .$$

Hence $\|T\| \le \max\{\|c_i\|_i,\ 0 \le i \le j\}$.

(ii) Let $W_2 = \big\{(a_0, a_1, \ldots, a_j)\big\}$, the a_i being elements of $L^i_s(H,F)$, the norm in W_2 being defined by: $\|a\| = \max\{\|a_i\|_i,\ 0 \le i \le j\}$. Then $c = (c_0, \ldots, c_j) \in W_2$. For fixed $i \in [0,j]$ define $w_i \in L^i_s(H, \mathbf{R}^1)$ of unit norm, and $w = (0, \ldots, w_i, 0, \ldots, 0) \in W_1$, where the $(j-i)^{\text{th}}$ component is w_i, the remaining components being 0's. Then $\|Tw\| = \|w_i \cdot c_{j-i}\|_j = \|c_{j-i}\|_{j-i} \le \|T\|\|w\| = \|T\|$. Hence for each $i = 0, 1, \ldots, j$, $\|c_{j-i}\|_{j-i} \le \|T\|$. Hence $\max_{0\le i\le j} \|c_i\|_i \le \|T\|$ i.e., $\|c\| \le \|T\|$.

From (i) and (ii) we conclude $\|T\| = \|c\|$. This proves the lemma.

This lemma implies that $\max_{0\le i\le j} \|D^i\big(\bar{g}(x) - g(x)\big)\|_i$ is the best constant $C > 0$ satisfying the inequality:

$$\text{(the sum in the left side of 8)} \quad \le C\Big[\|D^j h(x)\|_j + \binom{j}{1}\|D^{j-1}h(x)\|_{j-1} + \cdots h(x)\Big].$$

Hence

$$\max_{0\le i\le j} \|D^i\big(\bar{g}(x) - g(x)\big)\|_i \le \sigma(x) \, .$$

This is true for any $j = 0, 1, \ldots, k$. This is precisely what we wanted to satisfy.

Finally, putting together all the above conclusions (including that of Theorem A) we

find that for $j = 0, \ldots, k$,

$$\|D^j g(x) - D^j f(x)\|_j \leq \bar{\varepsilon}(x) + \sigma(x) = \frac{1}{2}\{\bar{\varepsilon}(x) + \varepsilon(x)\} < \varepsilon(x) .$$

This completes the proof of Theorem 9.

§10. Connection between strong approximation and earlier ideas of Bernstein-Nachbin

We can adopt a slightly different point of view concerning the space $C^k(\Omega, F)$, which will enable us to view the results of Kurzweil, or of Bonic and Frampton, or this author's results in [18], [19] as falling within the context of earlier ideas of Bernstein and L. Nachbin.

In order to explain this point of view we shall define

$$V = \{v(\cdot) = \frac{1}{\varepsilon(\cdot)} \mid \varepsilon(\cdot) \quad \text{is positive and continuous on} \quad \Omega\} ,$$

and denote by $C^k V_b(\Omega, F)$ the vector subspace of $C^k(\Omega, F)$ consisting of functions $f \in C^k(\Omega, F) \ni \max_{0<j\leq k} v(x)\|D^j f(x)\|$ is bounded on Ω for each $v \in V$. Each $v \in V$ determines a semi-norm by:

$$p_v(f) = \max_{0<j\leq k} \sup\{v(x)\|D^j f(x)\| \;\Big|\; x \in \Omega\}, f \in C^k(\Omega, F)$$

and we shall understand that $C^k V_b(\Omega, F)$ is endowed with the locally convex topology determined by the family of these semi-norms $\{p_v(\cdot)\}_{v \in V}$.

Now denote by W^ω the vector subspace of $C^0(\Omega, F)$ consisting of analytic mappings $\Omega \to F$ and by W^∞ the subspace of $C^k(\Omega, F)$ consisting of C^∞ mappings $\Omega \to F$. Then Kurzweil's theorem says that W^ω is dense in $C^0 V_b(\Omega, F)$. The theorem of Bonic and Frampton says that W^∞ is dense in $C^0 V_b(\Omega, F)$. Also the results of this author tell us that W^∞ is dense in $C^k V_b(\Omega, F)$ for a given integer $k \geq 0$.

§11. Strong approximation - other directions

The concept of approximation in the "Whitney topology" or "strong approximation" was seen before to be more suitable in the case of function spaces on non-compact spaces than that of uniform approximation or approximation in the compact open topology, and was not unrelated to the Bernstein-Nachbin ideas of weighted approximation. In a somewhat different direction, Whitney's idea of strong approximation has been modified in the realm of what have been called Nash mappings between Nash manifolds. In the case of Nash mappings of non-compact manifolds, a suitable "strong" topology is defined in the same way as the usual topology on the space $\mathcal{S}$ of rapidly decreasing C^∞ functions.

We shall give here a brief summary of some recent work of M. Shiota (cf. [55]) in this connection. This work of Shiota indicates some further possibilities Some preliminary definitions should first be introduced. Let $r = 0, 1, 2, \ldots,$ or ∞ or ω. A submanifold of $\mathbf{R}^n$ is called a *C^r-Nash manifold* if it is semi-algebraic and of class C^r; semi-algebraic (cf. [1]) means it is the finite union of sets defined by finitely many polynomial equalities and inequalities, i.e. a finite union of sets S where $x \in S$ if

$$P_1(x) = \ldots = P_j(x) = 0; \quad \text{and} \quad Q_1(x) > 0, \quad Q_2(x) > 0 \ldots, \quad Q_m(x) > 0,$$

the P's and Q's being polynomials in $x_1, \ldots, x_n$. A C^r-map from one C^r-Nash manifold to another is called a *C^r-Nash map* if the graph is semi-algebraic. A *C^r-Nash vectorfield* is similarly defined. $N^r(M)$ denotes the ring of all C^r-Nash functions on M where M is a C^r-Nash manifold. It is convenient and meaningful to restrict r to be $< \infty$, since it is known (see [55]) that a C^∞-Nash manifold and a C^∞-Nash map are already of class C^ω.

One objective of Shiota's paper is to approximate a C^r-Nash map between C^ω-Nash manifolds by a C^ω-Nash map. In the compact case an earlier result is attributed to Nash as also to Palais (see [55] for the references). In the non-compact case such a result is obtained in [55].

A stronger topology than either the uniform C^r-topology or the compact open topology, is introduced in $N^r(M)$, as follows. Let $f_k \in N^r(M), k = 1, 2, \ldots,$ and define "$f_k \to 0$ as $k \to \infty$" to mean the following: $v_1 \ldots v_{r'} f_k \to 0$ uniformly as $k \to \infty$, for any C^r-Nash vectorfields $v_1, \ldots, v_{r'}$ where $r' \leq r$. When $M = \mathbf{R}^n$ and $r = \infty$, this topology coincides with the usual topology on the space $\mathcal{S}$ of rapidly decreasing C^∞ functions (cf. Rudin [52]), i.e., $f_k \to 0$ as $k \to \infty \Leftrightarrow x^\alpha D^\beta f_k(x) \to 0$ uniformly, as $k \to \infty$, for any multi-indices α and β. The space $N^r(M)$ with this topology is *not* a vector space, because $af \not\to 0$ as $a \in \mathbf{R}$ converges to 0 unless supp (f) is compact. The strong topology defined above is called the *C^r-topology* on $N^r(M)$.

The following theorems proved in [55].

Theorem 11.1. *Let M_1 and M_2 be C^ω-Nash manifolds, and $f : M_1 \to M_2$ a C^r-Nash mapping. Then f can be approximated by a C^ω-Nash mapping in the C^r-topology. Further suppose the restriction of f to a given compact C^ω-Nash submanifold M_3 of M_1 to be of class C^ω. Then f can be approximated, fixing on M_3.*

The proof uses a C^ω-Nash function on $\mathbf{R}^n$ which is an approximation of 0 outside a small semi-algebraic neighbourhood of X (X being an algebraic set in $\mathbf{R}^n$), and of 1 in another one. This function is required to hold a useful well-known property of a C^∞-partition of unity. Several preliminary lemmas are needed for the proof, and though the

complete details of the proof of Theorem 11.1 would be outside the scope of this book, these lemmas themselves are interesting in their own right, and we shall explain these in some detail.

Let $f \in N^r(\mathbb{R}^n)$, and $e(x) = \frac{1}{C+|x|^{2k}}$, where C is a positive number, k is a positive integer, and $|x|^2 = x_1^2 + \cdots + x_n^2$. Write e as $e_{C,k}$ when it is necessary to emphasize C and k. Let $U \subset \mathbb{R}^n$ be an open semi-algebraic neighbourhood of $f^{-1}(0)$. Put

$$V_1 = \{x \notin U \mid f(x) > 0\}, \quad V_2 = \{x \notin U \mid f(x) < 0\} .$$

Lemma 11.2. *Define*

$$F = \frac{1}{2}\{(f^2 + e)^{\frac{1}{2}} + f\} .$$

Then $F \to 0$ on V_2 and $\to f$ on V_1 in the C^r-topology as C and $k \to \infty$ in such a way that $k^{2k} \leq C$.

Proof of Lemma 11.2. We shall suppose $r < \infty$. The problem is to prove $(f^2 + e)^{\frac{1}{2}} \to |f|$ on $V_1 \cup V_2$; hence it suffices to prove the convergence on V_1. The proof proceeds by induction on r.

Case $r = 0$. Let ε be a Nash function of the same form as e above. Put

$$\psi(t) = \inf\{f(x) \mid |x| = t \quad \text{and} \quad x \in V_1\} .$$

This function is positive and upper semi-continuous, and by the theorem of Tarski-Seidenberg, it is a semi-algebraic function on the closed semi-algebraic set

$$W = \{|x| \mid x \in V_1\} .$$

By Lojasiewicz's inequality (see Malgrange [40] p.59) and the stereographic projection, $\exists C_1, k_1 > 1 \ni$ for any $C \geq C_1$ and $k \geq k_1$,

$$\varepsilon(t)\psi(t) \geq \frac{1}{C + t^{2k}} \qquad \text{for} \qquad t \in W$$

where $\varepsilon(t)$ is defined such that $\varepsilon(|x|) = \varepsilon(x)$, i.e.,

$$\varepsilon(x) f(x) \geq e_{c,k}(x) \qquad \text{for} \qquad x \in V_1 \ .$$

Therefore

$$0 < (f^2 + e)^{\frac{1}{2}} - f = \frac{e}{(f^2 + e) + f} < \frac{e}{2f} \leq \frac{\varepsilon}{2} \qquad \text{on} \qquad V_1 \ .$$

Thus the proof for the case $r = 0$ follows.

Now suppose $(f^2 + e)^{\frac{1}{2}} \to f$ in the C^{r-1} topology; i.e., to be more precise, for any ε as above, $\exists C_2, k_2 \geq 1 \ni$ for any $C \geq C_2$, $k \geq k_2$ with $k^{2k} \leq C$ and a multi-index α with $|\alpha| \leq r - 1$, we have

$$\left| D^\alpha \{ f^2 + e_{C,k})^{\frac{1}{2}} \} - D^\alpha f \right| \leq \varepsilon \qquad \text{on} \qquad V_1 \ .$$

This inequality has to be proved with all $\alpha \ni |\alpha| = r$. Let α be one such multi-index. Then

$$\begin{aligned} D^\alpha e &= D^\alpha \Big\{ \{(f^2 + e)^{\frac{1}{2}} - f\}\{(f^2 + e)^{\frac{1}{2}} + f\} \Big\} \\ &= \Big\{ D^\alpha\big((f^2 + e)^{\frac{1}{2}} - f\big) \Big\} \cdot \Big\{ D^\alpha\big((f^2 + e)^{\frac{1}{2}} + f\big) \Big\} \\ &\quad + \sum_{\substack{\beta + \gamma = \alpha \\ \gamma \neq 0}} \Big\{ D^\beta\big((f^2 + e)^{\frac{1}{2}} - f\big) \Big\} \Big\{ D^\gamma\big((f^2 + e)^{\frac{1}{2}} + f\big) \Big\} \end{aligned}$$

Also $\exists$ constants $d_0, \ldots, d_{r-1}$ depending on r but not on C or on k, $\ni$

$$|D^\alpha e(x)| \leq \sum_{0 \leq i < r} d_i k^r |x|^{(r-i)(2k-1)-i} / (C + |x|^{2k})^{r-i+1} \ .$$

Now $k \leq k^{2k} \leq C \Rightarrow \frac{k^r|x|^{(r-i)(2k-1)-i}}{(C+|x|^2k)^{r-i}} \leq 1$. Thus

$$|D^\alpha e| \leq \sum_{i=0}^{r-1} d_i e \ .$$

This together with the induction hypothesis, implies

$$\left|D^\alpha\big((f^2+e)^{\frac{1}{2}} - f\big)\right| \leq \left\{de + \varepsilon_1 \sum_{0<\gamma\leq\alpha} \left|D^\alpha\big((f^2+e)^{\frac{1}{2}} + f\big)\right| / \big((f^2+e)^{\frac{1}{2}} + f\big)\right\}$$

on V_1, for any ε_1 as ε, and sufficiently small e with $k^{2k} \leq C$, where $d = \sum_{i=0}^{r-1} d_i$. Hence, as in the case $r = 0$, choosing small ε_1 we obtain $C_3, k_3 \geq 1 \ni \forall C \geq C_3$ and $\forall k \geq k_3$ with $k^{2k} \leq C$, we have

$$\left|D^\alpha\big((f^2+e)^{\frac{1}{2}} - f\big)\right| \leq \varepsilon \quad \text{on} \quad V_1 \ .$$

Thus the lemma follows.

Definition. *The argument for $r = 0$ in the preceding lemma is called Argument 1 (Arg. 1 in brief).*

Now let r' be a nonnegative integer $\ni r' < \infty$ and $r' \leq r$. Let ϕ be a polynomial on $\mathsf{R} \ni \phi(0) = \ldots = \phi^{r'}(0) = 0$. Then

$$\phi\{(|f| + f)/2\}$$

is a $C^{r'}$-Nash function which is r'-flat at $f^{-1}(0)$ i.e., $D^\alpha\phi\{(|f|+f)/2\} \equiv 0$ on $f^{-1}(0)$ for $|\alpha| \leq r'$.

Lemma 11.3. *$\phi(F) \to \phi\{(|f| + f)/2\}$ in the $C^{r'}$-topology as C and $k \to \infty$ so that $k^{2k} \leq C, F$ being as defined earlier.*

Proof of Lemma 11.3. Let $f_1 = (|f| + f)/2$, and α a multi-index with $|\alpha| \leq r'$. If $|\alpha| > 0$ then

$$D^\alpha \phi(F) = \phi'(F) D^\alpha F + \phi''(F) \sum_{\substack{\beta+\gamma=\alpha \\ \beta,\gamma\neq 0}} D^\beta F D^\gamma F + \cdots ,$$

and

$$D^\alpha \phi(f_1) = \begin{cases} 0 & \text{on} \quad f^{-1}(0)\,; \\ \phi'(f_1) D^\alpha f_1 + \phi''(f_1) \sum_{\substack{\beta+\gamma=\alpha \\ \beta,\gamma\neq 0}} D^\beta f_1 D^\gamma f_1 + \cdots & \text{outside } f^{-1}(0)\,. \end{cases}$$

Therefore, for any semi-algebraic neighbourhood U of $f^{-1}(0)$ the convergence $D^\alpha\phi(F) \to D^\alpha \phi(f_1)$ on $\mathbb{R}^n - U$ in the C^0 topology follows from Lemma 11.2. Hence it is enough to prove the following:

Suppose $\alpha_1, \ldots, \alpha_\ell > 0$ are multi-indices with $|\alpha_1| + |\alpha_2| + \cdots + |\alpha_\ell| = r'' \leq r'$. Let ε be a Nash function of the same form as e. Then $\exists C_1, k_1 \geq 1$ and $\exists$ an open semi-algebraic neighbourhood U of $f^{-1}(0) \ni \forall C \geq C_1$ and $\forall k \geq k_1$ with $k^{2k} < C$, we have

$$|\phi^\ell(F) D^{\alpha_1} F \ \ldots \ D^{\alpha_\ell} f_1| < \varepsilon \qquad \text{on} \qquad U\,,$$

$$|\phi^\ell(f_1) D^{\alpha_1} f_1 \ \ldots \ D^{\alpha_\ell} f_1| < \varepsilon \qquad \text{on} \qquad U - f^{-1}(0)\,.$$

Such a neighbourhood U satisfying the second inequality is readily seen to exist; thus consider the first one. By assumption, $\phi^{(\ell)}(F) = F^{r''-\ell+1}\psi(F)$ for some polynomial ψ. We have to show:

$$\left| F^{r''-\ell+1} D^{\alpha_1} F \ \ldots \ D^{\alpha_\ell} F \right| < \varepsilon \qquad \text{on} \qquad U\,,$$

which is equivalent to:

$$\left| f^{r''-\ell+1} D^{\alpha_1}(f_2 + f) \ \ldots \ D^{\alpha_\ell}(f_2 + f) \right| < \varepsilon \qquad \text{on} \qquad U\,,$$

where $f_2 = (f^2+e)^{\frac{1}{2}}$, because $\frac{f_2}{2} < |F| < f_2$. Furthermore, by induction on r'' (see below), and Argument 1, this last inequality is reduced to

$$\left|f_2^{r''-\ell+1} \dots f(x)\right| < \varepsilon \qquad \text{on} \qquad U\ .$$

Consider the case $r''=0$. Let

$$\begin{aligned} U &= \{x \in \mathbb{R}^n \;\Big|\; |f(x)| < \frac{\varepsilon(x)}{2}\}, \\ U' &= \{(x,t) \in U x \mathbb{R} \;\Big|\; (f^2(x)+t^2)^{1/2} < \varepsilon(x)\}\ . \end{aligned}$$

Then U' is an open semi-algebraic set containing $f^{-1}(0)$. By Argument 1, for arbitrarily small e, U' contains the graph of $e(x)^{1/2}$ on U; hence

$$\left(f^2(x)+e(x)\right)^{1/2} < \varepsilon(x) \qquad \text{on} \qquad U\ .$$

This proves Case $r''=0$.

Case $r''>0$. Invoking the argument for Case $r''=0$ and the equality

$$f_2^{r''-\ell+1} D^{\alpha_1} f_2 \ \dots\ D^{\alpha_\ell} f_2 \;=\; f_2 \prod_{i=1}^{\ell} f_2^{|\alpha_i|-1} D^{\alpha_i} f_2\ ,$$

we see that we should prove:

$$\left|f_2^{|\alpha_i|-1} D^{\alpha_i} f_2\right| \le C_2 + |x|^{2k_2}$$

for each i, some C_2, k_2 and arbitrarily small e with $k^{2k} \le C$. We note that

$$\left|D^{\alpha_i} f_2\right| \le C \sum_{\substack{\beta_1+\dots+\beta_k=\alpha_i \\ \beta_j>0}} \left|(f^2+e)^{\frac{1}{2}-k} D^{\beta_1}(f^2+e)\ \dots\ D^{\beta_k}(f^2+e)\right|$$

for some constant $C>0$. Hence we need to show

$$\left|(f^2+e)^{\frac{|\beta_j|}{2}-1} D^{\beta_j}(f^2+e)\right| < C_2 + |x|^{2k_2}\ .$$

This is clear if $|\beta_j| > 1$; if $|\beta_j| = 1$ then this follows from the inequality: $|D^{\beta_j} e| \leq d_0 e$ in the proof of Lemma 11.2. This completes the proof.

Next, given C_1, k_1, C_2, k_2 put $e_1 = e_{C_1,k_1}, e_2 = e_{C_1,k_2}$. Let $r' \ni r' < \infty$, and $r' \leq r$; let ϕ be a polynomial on $\mathbf{R} \ni \phi(0) = 0$, $\phi(1) = 1$, and $\ni$ if $r' \geq 1$, then $\phi' = \ldots = \phi^{(r')} = 0$ at 0 and 1. Put

$$\begin{aligned} F_1 &= \frac{1}{4}\Big(\big|3 - f - |f-1|\big| + 3 - f - |f-1|\Big), \\ F_2 &= \frac{1}{4}\left(\left\{\left(3 - f - \{(f-1)^2 + e\}^{\frac{1}{2}}\right)^2 + e_2\right\}^{\frac{1}{2}} + 3 - f - \left\{(f-1)^2 + e_1\right\}^{\frac{1}{2}}\right). \end{aligned}$$

Then $\phi(F_1)$ is a $C^{r'}$-Nash function $\ni \phi = 0$ on the set $\{f \geq 2\}$ and $= 1$ on the set $\{f \leq 1\}$. We now need the following lemma.

Lemma 11.4. *$\phi(F_1)$ can be approximated by $\phi(F_2)$ in the $C^{r'}$-topology by choosing e_1 and e_2 small.*

Proof. Fix e_1, temporarily. Set

$$F_3 = \frac{1}{4}\left\{\left|3 - f - \left((f-1)^2 + e_1\right)\right)^{\frac{1}{2}}\right| + 3 - f - \left((f-1)^2 + e_1)\right)^{\frac{1}{2}}\right\}.$$

Then $\phi(F_3)$ is a $C^{r'}$-Nash function. By Lemma 11.3, $\phi(F_2) \to \phi(F_3)$ in the $C^{r'}$-topology as C_2 and $k_2 \to \infty$ in such a way that $k_2^{2k_2} \leq C_2$. Hence we should show that $\phi(F_3) \to \phi(F_1)$ in the $C^{r'}$-topology as C_1 and $k_1 \to \infty$ satisfying $k_1^{2k_1} \leq C_1$.

The case $r' = 0$ follows from Lemma 11.3. Suppose $r' > 0$. We want an inequality: for small e_1

$$\Big|D^\alpha \phi(F_3) - D^\alpha \phi(F_1)\Big| \leq \varepsilon \tag{1}$$

where α is a multi-index with $0 < |\alpha| \leq r'$ and ε is a given function of the same form as

e. If $|\alpha| > 0$ then, as in proof of Lemma 11.3,

$$D^\alpha \phi(F_3) = \begin{cases} 0 & \text{on} \quad Y_3 \ , \\ \phi'(F_3) D^\alpha F_3 + \phi''(F_3) \sum_{\substack{\beta+\gamma=\alpha \\ \beta,\gamma\neq 0}} D^\beta F_3 D^\gamma F_3 + \cdots & \text{outside } Y_3 \end{cases}$$

$$D^\alpha \phi(F_1) = \begin{cases} 0 & \text{on} \quad Y_1 \cup Y_2 \ , \\ \phi'(F_1) D^\alpha F_1 + \phi''(F_1) \sum_{\substack{\beta+\gamma=\alpha \\ \beta,\gamma\neq 0}} D^\beta F_1 D^\gamma F_1 + \cdots & \text{outside } Y_1 \cup Y_2 \end{cases}$$

where $Y_1 = \{f = 1\}$, $Y_2 = \{f = 2\}$, and

$$\left\{3 = f + \left((f-1)^2 + e_1\right)^{\frac{1}{2}}\right\} = \{f = 2 - \frac{e_1}{4}\} \ .$$

By Argument 1, for sufficiently small e_1, Y_3 is contained in a given semi-algebraic neighbourhood of Y_2. Therefore to ensure (1) we only have to find open semi-algebraic neighbourhoods U_1 and U_2 of Y_1, Y_2 respectively, and C_{10}, k_{10} such that for each $\alpha_1, \ldots, \alpha_\ell > 0$ with $|\alpha_1| + \cdots + |\alpha_\ell| \le r'$ and any $C_1 > C_{10}$ and $k_1 > k_{10}$ with $k_1^{2k_1} \le C_1$, we have

$$\left|\phi^{(\ell)}(F_3) D^{\alpha_1} F_3 \ldots D^{\alpha_\ell} F_3\right| < \varepsilon \qquad \text{on} \qquad U_1 \cup U_2 - Y_3, \tag{2}$$

$$\left|\phi^{(\ell)}(F_1) D^{\alpha_1} F_1 \ldots D^{\alpha_\ell} F_1\right| < \varepsilon \qquad \text{on} \qquad U_1 \cup U_2 - Y_1 \cup Y_3, \tag{3}$$

$$\left|\phi^{(\ell)}(F_3) D^{\alpha_1} F_3 \ldots D^{\alpha_\ell} F_3 - \phi^{(\ell)}(F_1) D^{\alpha_1} F_1 \ldots D^{\alpha_\ell} F_1\right| < \varepsilon \qquad \text{on} \qquad \mathbb{R}^n - U_1 \cup U_2 \ .$$

Here F_1 and F_3 can be replaced by

$$F_{10} = (3 - f - |f-1|)/2$$

$$F_{30} = \left(3 - f - \left((f-1)^2 + e_1\right)^{\frac{1}{2}}\right)/2$$

respectively, because

$$F_1(x) = \begin{cases} F_{10}(x) & if \quad F_{10}(x) > 0 \\ 0 & otherwise \quad ; \end{cases}$$

$$F_3(x) = \begin{cases} F_{30}(x) & if \quad F_{3,0}(x) > 0 \\ 0 & otherwise \end{cases} ,$$

and furthermore,

$$F_{10}(x) > 0 \Leftrightarrow F_{30}(x) > 0 \quad \text{for} \quad x \in \mathsf{R}^n - U_1 \cup U_2 \text{and small} \quad e_1 \ .$$

We shall denote by $(2)_0, (3)_0$, and $(4)_0$ the respective replaced quantities.

(3_0) is trivial for some small U_1 and U_2 because

$$\phi^{(\ell)}(t) = t^{r'-\ell+1}(t-1)^{r'-\ell+1}\psi(t)$$

for some polynomial ψ. Then by Lemma 11.3, $F_{30} \to F_{10}$ on $\mathsf{R}^n - U_1$ in the $C^{r'}$-topology as C_1 and $k_1 \to \infty \ni k_1^{2k_1} \le C_1$ and this, together with $(3)_0$, proves $(2)_0$ on U_2 and $(4)_0$. Then as in Lemma 11.3, $(2)_0$ on U_1 is reduced to

$$\left| f^{r'-\ell+1} D^{\alpha_1} f_1 \ldots D^{\alpha_\ell} f_1 \right| < \varepsilon \quad \text{on} \quad U_1 \ ,$$

where $f_1 = \left((f-1)^2 + e\right)^{\frac{1}{2}}$. This is one of the inequalities in the proof of Lemma 11.3. Thus the Lemma follows.

Next let $X \subset \mathsf{R}^n$ be an algebraic set, I the ideal of $\mathsf{R}[x_1 \ \ldots, x_n]$ defined by X, namely consisting of polynomials vanishing on X, and h the square sum of finite generators of I. Define $f = h^{r'}/e_3$, with $C_3, k_3 > 1$ and $e_3 = e_{C_3,k_3}$. Define e_1, e_2, r', ϕ, F_1 and F_2 as we did earlier (before Lemma 11.4).

Lemma 11.5. *The functions $\phi(F_1)$ and $\phi(F_2)$ are $C^{r'}$, and C^ω Nash functions, respectively. Suppose U is a semi-algebraic neighbourhood of X. Then for small e_3, $\phi(F_1) = 0$ outside U and $= 1$ in another. Fix e_3. Then, for some e_1, and e_3, $\phi(F_2)$ is an approximation of $\phi(F_1)$ in the $C^{r'}$-topology.*

Proof. The first statement follows because of the definition. Next suppose e_3 is to chosen that $U \supset \{h < 2e_3\}$; this choice is possible by Argument 1. Then the second statement follows. The last statement is Lemma 11.4.

We shall denote by $Sing\,X$ the set of all the singular points of X. The next lemma follows.

Lemma 11.6. *Let $Y \subset X$ be a connected component of $X - Sing\,X$; let V be a semi-algebraic neighbourhood of $X - Y$ in $\mathbf{R}^n$. Let g be a C^r-Nash function on $\mathbf{R}^n$ r'-flat on Y. Then $g\phi(F_1) \to 0$ on $\mathbf{R}^n - V$ in the $C^{r'}$-topology as $C_3, k_3 \to \infty$ in such a way that $k_3^{2k_3} \le C_3$.*

Proof of Lemma 11.6. Suppose $\varepsilon \in N^\omega(R^b)$ be of the same form as e, and α a multi-index with $|\alpha| \le r'$. What requires to be shown now is:

$$\left|D^\alpha\big(g\phi(F_1)\big)\right| \le \varepsilon \qquad \text{on} \quad \mathbf{R}^n - V\ ,$$

for large C_3 and $k_3 \ni k_3^{2k_3} \le C_3$. This inequality can be reduced to simpler inequalities. Since $\phi(F_1) = 0$ outside $W = \{f \le 2\}$, it is enough to consider the inequality on $W - V$. Furthermore, since

$$\left|D^\alpha\big(g\phi(2-f)\big)\right| \ge \left|D^\alpha\big(g\phi(F_1)\big)\right| \qquad \text{everywhere,}$$

the earlier inequality can be replaced by

$$\left|D^\alpha\big(g\phi(2-f)\big)\right| \le \varepsilon \qquad \text{on} \quad W - V\ .$$

We note that if $|\gamma| > 0$ then

$$D^\alpha\big(g\phi(2-f)\big) \;=\; \sum_{\beta+\gamma=\alpha} D^\beta g D^\gamma \phi(2-f)$$

$$D^\gamma \phi(2-f) \;=\; \phi'(2-f)D^\gamma\phi(2-f) + \phi''(2-f) \sum_{\substack{\delta+\zeta=\gamma \\ \delta,\zeta\neq 0}} D^\delta(2-f)D^\zeta(2-f) + \cdots \,.$$

Then because $\phi'(2-f), \ldots, \phi^{(r')}(2-f)$ are bounded on W, therefore it is enough to prove:

$$\text{for} \quad |\beta| + |\alpha_1| + \cdots + |\alpha_\ell| \leq r', \qquad \left| D^\beta g D^{\alpha_1} f \,\ldots\, D^{\alpha_\ell} f \right| \leq \varepsilon \quad \text{on} \quad W - V \,.$$

This inequality in turn can be reduced to (for $|\beta| + |\alpha_1| + \cdots + |\alpha_\ell| + |\gamma_1| + \cdots + |\gamma_k| \leq r'$, and $H = h^{r'}$):

$$\left| D^\beta g D^{\alpha_1} H \ldots\; D^{\alpha_\ell} H D^{\gamma_1} e_3 \ldots D^{\gamma_k} e_3 / e_3^{k+1} \right| \;\leq\; \varepsilon$$

on $W - V$. If $\ell = 0$, then $k = -1$. In the proof of Lemma 11.2, the inequality: $\left| D^{\gamma_i} e_3^{k+1} \right| < \alpha e$ for some constant α, was established. Hence it is enough to prove:

$$\left| D^\beta g D^{\alpha_1} H \,\ldots\; D^{\alpha_\ell} H / H \right| \leq \varepsilon \quad \text{on} \quad W - V - Y \;,$$

where $|\beta| + |\alpha_1| + \cdots + |\alpha_\ell| \leq r'$ and $\ell \geq 1$, and

$$\left| D^\beta g \right| \leq \varepsilon \quad \text{on} \quad W - V$$

where $|\beta| \leq r'$. Here the inequality: $H \leq 2e_3$ on W as also the hypothesis: $D^\beta g = 0$ on Y, has been used.

Now consider the sets

$$Z \;=\; \left\{ x \in \mathbf{R}^n \;\middle|\; \left| D^\beta g D^{\gamma_1} H \,\ldots\; D^{\gamma_\ell} H(x) \right| \leq \varepsilon(x)\; H(x) \right\} \,,$$

$$Z' \;=\; \left\{ x \in \mathbf{R}^n \;\middle|\; \left| D^\beta g(x) \right| \leq \varepsilon(x) \right\} \,.$$

These sets are semi-algebraic and contain Y. Using Argument 1, it is enough to prove that Z and Z' are neighbourhoods of Y. For Z' this is clear. As regards Z, let $x_0 \in Y$

and consider a small neighbourhood of x_0. We can obtain a C^∞ local coordinate system $(y,z)=(y_1,\dots,y_{m_1}z_{m+1},\dots,z_{n\cdot})$ around $x_0 \ni (y,z)=0$ at x_0 and

$$h(y,z) \;=\; y_1^2+\cdots+y_m^2 \qquad \text{and} \quad Y \;=\; \{y_1=\cdots=y_m=0\}$$

(cf. Lemma 4.11 in Shiota [55]). Then by hypothesis

$$\left|D_x^\beta g(y,z)\right| \;\le\; d'\left|y\right|^{r'+1-|\beta|}$$

in a neighbourhood of 0 for some constant d'. Hence

$$\begin{aligned}\left|D_x^\beta g D_x^{\alpha_1}H\ \dots\ D_x^{\alpha_l}H(y,z)\right| &\le r''|u|^{\gamma+1-|\beta|+2r'-|\alpha_1|+\cdots+2r'-|\alpha_l|}\\ &\le d''|y|^{2r'+1}\ .\end{aligned}$$

in a neighbourhood of 0 for some constant d''. Hence Z contains a neighbourhood of x_0. The lemma follows.

We shall only state the next lemma. However to state this lemma it is necessary to explain some further notation and definitions. Let $h \in N^\omega(\mathbf{R}^n)$, $X=h^{-1}(0)$, $U \subset X$ a connected C^ω-Nash manifold open in X and $g \in N^\omega(U)$. A *minimal polynomial* $P(z,x)$ *for* g means a polynomial in $n+1$ variables $\ni P(z,x)\Big|_{\mathbf{R}\times U} \not\equiv 0$, $P\big(g(x),x\big) \equiv 0$ on U, and the degree in z is minimal. Then we say that the *pair* (g,P) *has Property* $(A1)$ if $P^{-1}(0)\cap(\partial P^{-1}/\partial z)(0)\cap\mathbf{R}\times U=\emptyset$, P is of constant degree in z at every point of U and $\{P^{-1}(0)\cup(\partial P^{-1}/\partial z)(0)\}\cap\mathbf{R}\times U$ is the disjoint union of the graphs of C^ω-Nash functions on U.

By induction, if the pair of each C^ω-Nash function on U whose graph is contained in $(\partial P^{-1}/\partial z)(0)$ and some minimal polynomial for it has property $(Ak-1)$ then we say (g,P) *has property* (Ak) *for* $k>1$. *Property (A)* is simply property (Am) for m, this being the degree of P *in* z.

Suppose (g, P) has Property (A). Then we say (g, P) is *of height* 0 if $(\partial P^{-1}/\partial z)(0) \cap \mathbf{R} \times U = \emptyset$ and, inductively, (g, P) is of *height* ℓ if it is *not* of height $\ell - 1$ and the pair of each C^ω-Nash function on U defined by $(\partial P^{-1}/\partial z)(0)$ and some minimal polynomial for it is of height $\leq \ell - 1$. If (g, P) has Property $(A1)$ then g can be extended uniquely to some semi-algebraic neighbourhood of U in $\mathbf{R}^n$ satisfying $P\big(g(x), x\big) \equiv 0$. The extension will be written as $\tilde{g}_P$; the domain of $\tilde{g}_P$ will not be specified.

The next lemma then is as follows.

Lemma 11.7. *Let $D \subset \mathbb{R}^n$ be a closed semi-algebraic set contained in U. Suppose the pair (g, P) where $g \in N^\omega(U)$ and P is a polynomial, has property (A). Then $\exists$ closed semi-algebraic neighbourhood $\tilde{D}$ of D in $\mathbb{R}^n \ni \tilde{g}_P$ is defined on $\tilde{D}$ and $\tilde{g}_P\big|_{\tilde{D}}$ can be approximated in the C^ω-topology by the restriction to $\tilde{D}$ of a C^ω-Nash function on $\mathbb{R}^n$.*

We shall leave it to the reader to read the further details of the proof in [55].

This Page Intentionally Left Blank

CHAPTER IV

Approximation problems in probability

We shall turn now to some approximation results in the realm of probability. Again, as in the earlier three chapters, we have confined ourselves to only a few aspects of this topic, for several reasons, viz. because this aspect of probabilistic approximations fitted in with the topics dealth with earlier in this monograph, and also because several other aspects have been exhaustively dealt with by other authors. For example, there are beautiful expositions of the topics of convergence of distributions and the classical theorems of P. Levy, Khinchin, Prokhorov, Berry-Essen and Kolmogorov, (cf. Gnedenko and Kolmogorov [14], and M. Rosenblatt [50]). Further the topic of weak convergence has been dealt with thoroughly by various authors such as D. Pollard [48] and others (cf. Ito and McKean [22], and Ito and Nisio [23]).

As is well-known by now, probability provides an interesting and extremely useful tool in analysis as also in applied problems. We shall begin this chapter with an interesting and elementary application of probabilistic concepts to obtain a proof of the familiar Weierstrass' theorem on uniform approximation of a continuous function on a compact interval by a polynomial. To be specific we shall present S. Bernstein's proof of this theorem (cf. [4]) in which he uses his "Bernstein's Polynomials" (though this is well-known, cf. [13], [51]).

§1. Bernstein's proof of Weierstrass' theorem

It is enough to consider a real-valued function $f(x)$ defined and continuous on the unit interval $[0,1]$. Let Y be a binomial random variable (cf. Appendix 4) of sample size n, that is to say, corresponding to n coin tosses, x, being taken to be the probability of one success. We then consider $f(Y/n)$ as one estimate of $f(x)$. This estimate equals $f(k/n)$ for $k = 0, 1, \ldots, n$, with probability $\binom{n}{k}x^k(1-x)^{n-k}$. The *Bernstein polynomial* $p_n(x)$ *of degree* n in x is defined by:

$$p_n(x) = Ef(Y/n) = \sum_{k=0}^{n} f(k/n)\binom{n}{k}x^k(1-x)^{n-k},$$

this being the "expectation", or the mean value of the variable $f(Y/n)$. The objective then is to show that $p_n(x)$ converges uniformly to $f(x)$ as $n \to \infty$.

Let $\varepsilon > 0$. Then uniform continuity of f on $[0,1]$ implies that $\exists\delta(\varepsilon) > 0 \ni \forall x, y \in [0,1]$ with $|x-y| < \delta(\varepsilon)$, we have $|f(x) - f(y) < \varepsilon$. Now we apply Chebyehev's inequality (cf. Appendix 4), and obtain

$$\begin{aligned} P\left[\left|\frac{Y}{n} - x\right| \geq \eta\right] &\leq \sigma^2\left(\frac{Y}{n}\right)/\eta^2 = \frac{x(1-x)}{n\eta^2}, \\ &\leq \frac{1}{4n\eta^2} \end{aligned}$$

σ^2 being the variance of the binomial variable Y. Now set $\eta = \frac{1}{2}\delta(\frac{\varepsilon}{2})$, and $M = \max\{|f(x)| \mid 0 \leq x \leq 1\}$. Then if $n > \frac{2M}{\varepsilon\delta^2(\varepsilon/2)}$, we find

$$\begin{aligned} |p_n(x) - f(x)| = \left|Ef\left(\frac{Y}{n}\right) - f(x)\right| &\leq MP\left[\left|\frac{Y}{n} - x\right| \geq \frac{1}{2}\delta\left(\frac{\varepsilon}{2}\right)\right] + \frac{\varepsilon}{2} \\ &\leq \frac{M}{n\delta^2(\varepsilon/2)} + \frac{\varepsilon}{2} \leq \varepsilon\,. \end{aligned}$$

This completes Bernstein's proof.

More recently, Bernstein's idea of using probabilistic techniques in approximation problems in analysis has been pushed further. We shall present a few of these recent results. However before turning to these recent developments we wish to note that Bernstein's polynomials have been further used, to obtain accurate estimates of the errors of approximation, and these results have proved useful in semigroup theory (cf. Butzer and Behrens [6]).

§2. Some recent Bernstein-type approximation results

In the preceding section we saw that for a function $f \in C[0,1]$, the Bernstein polynomials

$$B_n f(x) = \sum_{k=0}^{n} f(\frac{k}{n}) p_{n,k}(x)$$

where $p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$ converge to $f(x)$ uniformly on $[0,1]$. Here $B_n f(x)$ can be regarded as the result of an operator B_n operating on $f \in C[0,1]$, for $n = 1, 2, \ldots$. In the same context it is known that

$$\Big| B_n(f;x) - f(x) \Big| \le \frac{5}{4} \omega(n^{-1/2}) ,$$

where $\omega(\delta) = \sup\{|f(x) - f(y)| \;\Big|\; |x-y| \le \delta, \quad 0 \le x, y \le 1\}$, $\delta > 0$ (Popoviciu, cf. [38] p. 20). Further suppose $f^{(2k)}(x)$ exists at x, then (Bernstein, cf. [38] pp. 22-23)

$$\begin{aligned} \lim_{n\to\infty} n^k \left[B_n(f;x) - f(x) - \left[\sum_{s=1}^{2k-1} \frac{n^{-s}}{s!} T_{ns}(x) f^{(s)}(x) \right] \right. \\ = \left(\frac{x(1-x)}{2} \right)^k \frac{f^{2k}(x)}{k!} \end{aligned}$$

where $T_{ns}(x) = \sum_{k=0}^{n} (k - nx)^s p_{n,k}(x)$, $n = 1, 2, \ldots$, $s = 0, 1, 2, \ldots$.

A number of such operators have been introduced after Bernstein. Many of these are special cases of an operator introduced by Feller [13].

Feller's operator is defined as follows. Let $\{X_n, n \ge 1\}$ be a sequence of r.v.'s with distribution function $(d \cdot f.)$ $F^*_{n,x}(t)$ with expectation $EX_n = x$, and variance $\sigma_n^2(x)$, x being a continuous real parameter. For a continuous function f on $\mathbb{R}^1$ define

$$L_n f(x) = Ef(X_n) = \int_{-\infty}^{\infty} f(t) dF^*_{n,x}(t) \qquad \text{if} \quad E|f(X_n)| < \infty .$$

The following lemma of Feller [13] is known.

Lemma 2.1. *If $\sigma_n^2(x) \to 0$ as $n \to \infty$, then $\lim_{n\to\infty} L_n f(x) = f(x)$ for every continuous bounded function f. If further f is uniformly continuous and $\sigma_n^2(x) \to 0$ uniformly, w.r.t. $x \in \mathbb{R}^1$, then the convergence of $L_n f(x)$ to $f(x)$ is uniform.*

Now suppose the continuous parameter x takes values in an interval I (perhaps infinite); let $G(x)$ be a *d.f.* on I. We then obtain:

Lemma 2.2. *Suppose $\sigma_n^2(x) \leq g(x)$, where $g(\cdot)$ is G-integrable and that the conditions of Lemma 1 are valid. Then*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} L_n(f;x)dG(x) = \int_{-\infty}^{\infty} f(x)dG(x) .$$

The preceding scheme is now modified as follows. Let $\{Y_n, n \geq 1\}$ be a sequence of i.i.d. (independent, identically distributed) r.v.'s with mean $x \in I$ and variance $\sigma^2(x)$. Let $S_n = \sum_{i=1}^{n} Y_i$. Then the above expression for $L_n f(x)$ is modified to the following:

$$L_n f(x) = Ef(S_n/n) = \int_{-\infty}^{\infty} f(t/n) dF_{n,x}(t), \tag{1}$$

where $F_{n,x}(t)$ is the distribution function of S_n.

The following theorem has been obtained by R.A. Khan (cf. [28] p. 195), thus extending, the above-mentioned result of Popoviciu and Bernstein to Feller's operator defined in (1).

Theorem 2.3. *Suppose $\{Y_n, n \geq 1\}$ is a sequence of i.i.d. r.v.'s with mean $x \in I \subset \mathbb{R}^1$, and variance $\sigma^2(x)$. Let $A = \sup_{x\in I} \sigma^2(x)$. Then for $x \in I$ the Feller operator defined in (1) above satisfies:*

$$\left|L_n(f,x) - f(x))\right| \leq (1 + A)\omega(f; n^{-1/2})$$

where $\omega(f;\delta) = \sup\{|f(x)-f(y)| \,\big|\, |x-y| \le \delta, x, y \in \mathbf{R}^1\}$. *Furthermore, for* $x \in [\alpha,\beta] \subset I$,

$$|L_n(f,x) - f(x)| \le \{1 + \sup_{\alpha \le x \le \beta} \sigma^2(x)\} \cdot \omega(f; n^{-1/2}) .$$

Proof of Theorem 2.3. Let $S_n = \sum_{i=1}^{n} Y_i$ and $\lambda = \lambda\left(\frac{S_n}{n}\right) = \left[\frac{1}{\sigma}\left|\frac{S_n}{n} - x\right|\right]$ where $[r]$ denotes the greatest integer $\le r$. Clearly $\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \le \omega(f;\delta)(1+\lambda)$, and hence

$$\begin{aligned} \left|L_n(f;x) - f(x)\right| &\le \omega(f;\delta)E(1+\lambda) \le \omega(f;\delta)(1+E\lambda^2) \\ &\le \left\{1 + \frac{E(S_n - nx)^2}{n^2\delta^2}\right\}\omega(f;\delta) \\ &= \left\{1 + \frac{\sigma^2(x)}{n\delta^2}\right\}\omega(f;\delta) . \end{aligned}$$

Now let $\delta = n^{-1/2}$, and the proof is completed.

As regard monotonic convergence, the following theorems are proved in [26] p. 199 for the Feller operator defined earlier, when the function $f(x)$ is convex.

Theorem 2.4. *Let* $\{Y_n, n \ge 1\}$ *be a sequence of i.i.d. r.v.'s with mean* $x \in I$ *and variance* $\sigma^2(x)$. *For a continuous convex and bounded function* f *on* $\mathbf{R}^1$ *defined the Feller operator by:*

$$L_n(f;x) \;=\; Ef(S_n/n) \;=\; \int_{-\infty}^{\infty} f(t/n)dF_{n,x}(t) \;,$$

where $F_{n,x}(t)$ *is the d.f. of* $S_n = \sum_{i=1}^{n} Y_i$. *Then* $L_n(f,x) \ge L_{n+1}(f,x) \ge \ldots \ge f(x)$, *and* $L_n f(x) \searrow f(x)$ *uniformly on every bounded interval.*

For the proof, the following lemma is needed.

Lemma 2.5. *Let* $Y_1, Y_2 \ldots$ *be i.i.d r.v.'s with finite expectation. Let* $S_n = \sum_{i=1}^{n} Y_i$. *Then*

$$E\left\{\frac{S_n}{n} \,\Big|\, S_{n+1}\right\} \;=\; \frac{S_{n+1}}{n+1} \quad \text{a.e.}$$

Proof of Lemma 2.5. We note that $E(Y_1 \big| S_{n+1}) = E(Y_2 \big| S_{n+1}) = \ldots = E(Y_{n+1} \big| S_{n+1})$; hence $E(Y_i \mid S_{n+1}) = E(\frac{\sum_{i=1}^{n} Y_i}{n+1} \mid S_{n+1}) = \frac{S_{n+1}}{n+1}$. Thus

$$E\Big(\frac{S_n}{n} \mid S_{n+1}\Big) = \frac{1}{n} \sum_{i=1}^{n} E\Big(\frac{Y_i}{S_{n+1}}\Big) = \frac{S_{n+1}}{n+1} \quad \text{a.e.}$$

This proves the Lemma.

Proof of Theorem 2.4. We note that

$$L_n(f;x) = Ef\Big(\frac{S_n}{n}\Big) = E\left(E\Big(f\Big(\frac{S_n}{n}\Big) \mid S_{n+1}\Big)\right) .$$

The function f is assumed to be convex; hence using the conditional version of Jensen's inequality and the preceding Lemma, we find

$$\begin{aligned} L_n(f;x) &\geq Ef\left(E\Big(\frac{S_n}{n} \mid S_{n+1}\Big)\right) = Ef\Big(\frac{S_{n+1}}{n+1}\Big) \\ &= L_{n+1}(f;x) . \end{aligned}$$

Thus

$$L_n(f;x) \geq L_{n+1}(f;x) \geq \ldots \geq f(x) .$$

Now using the very first lemma in this topic (Lemma 2.1), the proof of the theorem is completed.

More such results concerning Bernstein type operators are due, e.g., to Katherine Balazs (cf [3]). Some of the results in [3] are as follows. However, these are not exactly probabilistic in nature!

For a function f on the positive half axis, with a_n, b_n positive numbers $\ni b_n \to \infty$,

and $a_n = \frac{b_n}{n} \to 0$ as $n \to \infty$, define the Bernstein type rational functions $R_n(f;x)$ by

$$R_n(f;x) = \frac{1}{(1+a_n x)^n} \sum_{k=0}^{n} f\Big(\frac{k}{b_n}\Big)\binom{n}{k}(a_n x)^k$$
$$\sum_{k=0}^{n} f\Big(\frac{k}{b}\Big) r_{kn}(x); \qquad x \geq 0, \quad n = 1, 2, \dots .$$

This particular positive linear operator has been investigated (see the references in [3]).

For a function f defined in $(-\infty, \infty)$, define the nth Bernstein type rational function $R_n^*(f;x)$ by:

$$R_n^*(f;x) = \frac{1}{(1+a_n x)^n + (1-a_n x)^n} \sum_{k=0}^{n} \Big\{f\Big(\frac{k}{b_n}\Big) + (-1)^k f\Big(-\frac{k}{b_n}\Big)\Big\}\binom{n}{k}(a_n x)^k$$
$$= \sum_{k=0}^{n} \Big\{f\Big(\frac{k}{b_n}\Big) + (-1)^k f\Big(-\frac{k}{b_n}\Big)\Big\} r_{kn}^*(x), \qquad -\infty < x < \infty ,$$

where $n > 0$ is *even*, and a_n, b_n satisfying: a_n, b_n are > 0, $b_n \to \infty$, $a_n = \frac{b_n}{n} \to 0$, as $n \to \infty$. The theorem established is

Theorem 2.6. *(cf. Theorem 2 [3], p. 196). Suppose f is continuous in $(-\infty, \infty)$ and satisfies: $f(x) = O(e^{\alpha|x|})$ for some $\alpha > 0$. Then for arbitrary fixed $A > 0$ and $\varepsilon > 0$,*

$$|f(x) - R_n^*(f;x)| \leq c_1 \omega_{[-A-\varepsilon, A+\varepsilon]}\big(f; \max\{a_n, b_n^{-1/2}\}\big)$$

for $-A \leq x \leq A$; here $n > 0$ is even, $\omega_{[-A-\varepsilon,A+\varepsilon]}(f;\cdot)$ denotes the modulus of continuity of f in $[-A-\varepsilon; A+\varepsilon]$, and $c_1 = c_1(\alpha; A; \varepsilon) > 0$ is a number independent of n.

Furthermore, a necessary and sufficient condition for the uniform convergence of $R_n^*(f)$ to f is: $f \in C[-\infty, \infty]$ where where $C[-\infty, \infty]$ denotes the class of continuous functions $f \ni \lim_{|x|\to\infty} f(x)$ exists (finite) and hence f is uniformly continuous on $[-\infty, \infty]$. The precise theorem is:

Theorem 2.7. *(cf. Theorem 4, [3], p. 197). Suppose $b_n = n^\beta$ with $0 < \beta \le \frac{2}{3}$, Then*

$$\lim_{n\to\infty} \sup_{-\infty<x<\infty} \left|f(x) - R_n^*(f;x)\right| = 0, \qquad n > 0 \text{ even} \quad \Leftrightarrow f \in C[-\infty,\infty] .$$

A slightly more abstract situation is considered by R. Wittman (cf. [64]). Suppose E is a locally convex space, and $A \subset E$ a convex set. A function $f : A \to \mathbb{R}^1$ is uniformly continuous if $\exists$ continuous seminorm p on $E \ni$

$$|f(x) - f(y)| \le p(x - y) \qquad \forall x, y \in A .$$

The result established in the following:

Theorem 2.8. *(cf. Theorem 1, [64], p. 463). Suppose $f : A \to \mathbb{R}^1$ is a uniformly continuous convex function. Then for every $\varepsilon > 0$ $\exists$ Lipschitz continuous convex function g on E satisfying:* $\sup\limits_{x\in A} |g(x) - f(x)| \le \varepsilon$.

§3. A theorem of Steinhaus

Probabilistic ideas have been used in a function theoretic context, specifically in the context of random Taylor series. E. Borel, in 1896, formulated the statement that, with probability one, the circle of convergence of a power series with arbitrary coefficients is its natural boundary, i.e. consists only of singular points. This statement is only about plausibility. The correct formulation was given by H. Steinhaus [59] in 1929. The result can be stated as follows.

Theorem 3.1. *The series* $\sum\limits_{0}^{\infty} r_n e^{i\phi_n} z^n$, *where the* ϕ_n*'s are mutually independent random variables uniformly distributed on* $[0, 2\pi]$, *and* $\limsup\limits_{n\to\infty} r_n^{1/n} < \infty$, *has almost surely the circle of convergence as its natural boundary.*

We shall give here an account of a theorem of H. Steinhaus in this context, formulated for Steinhaus series. This account follows Kahane [26], where also one can find references to further related developments. At this point a brief explanation of the probabilistic ideas involved in this theorem, appears to be in order, and the explanation given in the next paragraph appears to be the bare minimum of "probability theory" needed for this theorem. A more general account will be found in Appendix 4.

Probabilistic ideas; random variables

A random point is a point $\omega = (\omega_1, \omega_2, \ldots)$ in the hypercube $\mathcal{H} = [0,1] \times [0,1] \times \cdots$ of infinite dimensions. An event is a subset $\mathcal{E} \subset \mathcal{H}$ and its probability $p(\mathcal{E})$ is its (infinite dimensional) product Lebesgue measure, if it exists. If $p(\mathcal{E}) = 1$ we say that $\mathcal{E}$ occurs *almost always* (a.a.), of that $\mathcal{E}$ is *almost certain*. A *random variable* (r.v.) is a measurable function of ω, and is always denoted by a capital letter. With any hypothetical property

of a random variable or of a family of random variables we speak of the event of the occurrence of this property; we thus speak of the property occurring almost always. The *expected value* (or *expectation*) of a r.v. T is $E(T) = \int_{\mathcal{H}} T(\omega)d\omega$. For any given sequence $\{T_k\}$ of r.v.'s if T_k depends only on the kth component ω_k, we have $E(\prod_k T_k) = \prod_k E(T_k)$. The *characteristic function* of a r.v. T is $E(e^{iuT})$ where $u \in \mathbb{R}^1$; if $E(e^{iuT}) = e^{-u^2/2}$, T is said to be a (real) *normal r.v.*. Two r.v.'s T_1 and T_2 are *orthogonal* if $E(T_1T_2) = 0$.

The r.v.'s $T_j (j = 1, 2, \dots,)$ are said to be *independent* if $\exists$ a mapping $\omega \to \omega'$ of $\mathcal{H} \to \mathcal{H}$, which is measure-preserving $\ni$ under this mapping T_j = function of ω'_j a.a. Two r.v.'s T, T' are said to follow (or to be *subject to*) the *same law* if $T'(\omega) = T(\omega')$ a.a. A complex r.v. T is *invariant under rotation* if for each $t \in \mathbb{R}^1$, T, and Te^{it} are subject to the same law. If the r.v.'s $T_j, j = 1, 2$, are real normal and independent, then for each $t \in \mathbb{R}^1$, $T_1 \cos t + T_2 \sin t$ is normal, and $T_1 + iT_2$ is called a *complex normal* r.v. A complex normal r.v. is invariant under rotation. If T is a normal r.v. (real or complex), then $a + bT$ is a *Laplacian* (or *Gaussian*) r.v., a, b being constants.

We shall need the following lemmas; for the proofs of the first two cf. Loeve [37], and for the third cf. Zygmund [67].

Kolmogorov's Lemma. Given an infinite sequence $\{T_j\}$ of independent r.v.'s, and an event $\mathcal{E}$ which for each $n = 1, 2, \dots,$ does not depend upon the realisation (i.e. on the values of) T_j, $j = 1, 2, \dots, n$, then $p(\mathcal{E}) = 0$ or 1.

Khinchin's Lemma. If the T_j are real independent r.v.'s, $\ni$ $\mathsf{E}(T_j) = 0$, $E(T_j^2) = 1$ for $j = 1, 2, \dots,$ then the series $\sum_j c_jT_j$ converges a.a. provided $\sum |c_j|^2 < \infty$.

Zygmund's Lemma. If the r.v.'s $T_j, j = 1, 2, \dots,$ are subject to the same law,

then the series $\sum_j c_j T_j$ is not summable under any regular process of summability if $\sum_j |c_j|^2 = \infty$.

Random trigonometric series

Here we are concerned with random functions ω into $F_\omega(t)$ mapping $\mathcal{H}$ into a function on the unit circle. Consider a local property of F, e.g. $F(t)$ being > 0; or $F(t)$ being continuous in t etc. We say at almost all points t, the property holds almost certainly meaning: the property holds almost with probability 1 at all points t except perhaps for t lying in a set of measure 0 on the unit circle. If this happens we say the property holds almost certainly (surely) almost everywhere, (a.s.a.a.). If a property holds at each t almost certainly, it does not follow that it holds almost certainly for all t. It is often much more difficult to determine the probability that a local property should hold everywhere than it is to determine such a probability at any (fixed) point.

A r.v. is *(strictly) stationary* if the probabilities associated with it are invariant under translations $t \to t - t_0$, that is to say, if to each t_0 $\exists$ transformation $\omega \to \omega'$ (see earlier in this section) $\ni$ $F_{\omega'}(t) = F_\omega(t - t_0)$. To say that a local property should hold a.s., a.e. is saying that is should hold a.s. at a fixed point t_0. The convolution of a stationary r.v. with a function which is certain (i.e. not subject to any randomness) is a r.v. which is stationary.

It is also quite natural to consider formal trigonometric series $\sum A_n e^{int}$, with coefficients A_n which are complex r.v.'s. Such a series is said to be *stationary* if its convolution with anay trigonometric polynomial is stationary.

We shall consider here only *Steinhaus series* F in which the coefficients A_n are in-

dependent r.v.'s which are invariant under rotations. We shall use the notation $F \sim \Sigma A_n e^{int}$, so as to be able to use this symbol to identify F with a function, or a distribution, as the case may be.

Steinhaus' theorem

We consider properties P of trigonometric series f on intervals, which are subject to the following conditions:

(1°) If f satisfies P on (a, b), any translate f_t satisfies P on $(a + t, b + t)$.

(2°) If f satisfies P on two abutting intervals, then f satisfies P on their union.

(3°) If f satisfies P on an interval, then so does $f + p$, where p is any trigonometric polynomial.

(4°) If F is a Steinhaus series, then the statement: "F satisfies P on (a, b)" is an event for which the probability exists.

The theorem of Steinhaus referred to above is as follows.

Theorem 3.1. *(cf. [26]). Consider a Steinhaus series F and a property P. Then almost surely, F satisfies P everywhere or F satisfies P nowhere.*

Proof. Let $\{I_n\}$ be a finite collection of intervals, all of the same length ε, which cover the circle, and

$$x_{\varepsilon,n} = p[\text{on} \quad I_n, \ F \quad \text{satisfies} \quad P], \qquad n = 1, 2, \ldots,$$

and

$$x_\varepsilon = p[F \quad \text{satisfies } P \text{ on at least one} \quad I_n] \ .$$

The condition (4°) in the definition of the property P implies that $x_{\epsilon,n}$ and x_ϵ both exist. Suppose $x_\epsilon > 0$. Now $x_\epsilon \leq \sum_n x_{\epsilon,n}$, condition (1°) implies that, since F is stationary, all the $x_{\epsilon,n}$ are equal, hence $x_{\epsilon n} > 0$. Condition (3°) and the independence of the A_n's implies that $x_{\epsilon,n} = 1$ (by Kolmogorov's Lemma). Then condition (2°) implies that $p[F$ satisfies P everywhere$] = 1$.

The following corollaries are noteworthy.

Corollary 3.2. *Every Steinhaus-Taylor series has its circle of convergence (if it exists) as its natural boundary.*

Corollary 3.3. *If a Fourier-Steinhaus series represents a.s. a distribution F, and if with positive probability $\exists$ interval (dependent on ω) on which F equals a continuous (resp. analytic) function, then a.s. F is a continuous (resp. analytic) function, etc.*

As a special case of the series dealt with in Steinhaus' theorem, we would like to mention *Brownian motion*, (or Wiener measure) viz. the series $\sum_{n=\infty}^{\infty} \frac{1}{n} Z_n e^{nit}$, where $Z_n = X_n + iY_n$, X_n, Y_n being independent real normal r.v.'s. The series is continuous a.s. and a.s. satisfies a Hölder condition of exponent $< \frac{1}{2}$, but a.s. does not satisfy a Hölder condition of exponent $> \frac{1}{2}$. It is also known that up to normalisation Brownian motion is the only process $\{X_t\}$, $0 \leq t \leq 2\pi$ with a.s. continuous trajectories such that the X_t are normal and $X_t - X_s$ independent over disjoint intervals (cf. L. Schwartz [50]).

However, in the next section we prefer to sketch a slightly different derivation of Browman motion using the Haar functions.

§4. The Wiener process or Browman motion

First we shall state the definition of the Wiener process. In the following T shall denote the closed interval $[0,1]$ or $\mathbb{R}_+ = [0,\infty)$. The reader should refer to the Appendix 4 for the meaning of the standard probabilistic terminology used here in this section.

Definition. A process $X = (X_t)$ defined on a complete probability space $(\Omega, \mathcal{A}, P)$ is called a *Wiener process with variance parameter* σ^2 if it is a Gaussian process with the properties:

1. $X_0(\omega) = 0$ (a.s.);
2. $\forall s,t$ with $s \le t$, $X_t - X_s$ has a Gaussian distribution with zero mean and variance $\sigma^2(t-s)$;
3. $\forall t_i \in T(i = 1,2,3,4) \ni t_1 \le t_2 \le t_3 \le t_4$, the r.v.'s $X_{t_2} - X_{t_1}$, $X_{t_4} - X_{t_3}$ are independent,
4. for a.a. ω, the trajectories $t \to X_t(\omega)$ are continuous.

An equivalent definition is as follows; it happens to be a little more convenient in certain circumstances.

Definition. $X = (X_t)$ is a *Wiener process* with variance parameter σ^2 if X is a continuous Gaussian process with $EX_t = 0$ $\forall t$ and covariance function given by

$$E(X_t X_s) = \sigma^2 \min(t,s) .$$

A Wiener process with $\sigma^2 = 1$ is called a *standard Wiener* process (or *standard Brownian motion*).

First let $T = [0,1]$. Several methods of constructing a Wiener process on T are known (cf. Ito and McKean [22]). The particular method sketched below is based on the

use of the Haar family of functions on $[0,1]$ (cf. Cisielsky [7], P. Levy [35]). However, since the excellent and complete exposition of this method in Kallianpur's monograph [27] seems impossible to improve upon, we shall only sketch the highlights of the proof.

First we note that a $N(a,\sigma^2)$ r.v. ξ induces a Borel probability measure $\mathbb{P}$ on $\mathbb{R}$, by the rule: $\mathbb{P}(A) = P\{\xi^{-1}(A)\}$, A a Borel set in $\mathbb{R}$. Now let $\Omega = \mathbb{R}^\infty$, the countable product of real lines, $B(\mathbb{R}^\infty)$ the σ-field generated by the Borel-cylinder sets in Ω, and $\mathbb{P}$ the countable product of $N(0,1)$ measure on $\mathbb{R}$. Denote by $\mathcal{A}$ the completion of $B(\mathbb{R}^\infty)$ with respect to $\mathbb{P}$.

We then introduce the Haar system of functions $\{g_{n,j}\}$ on $[0,1]$, where $n = 1, 2, \ldots,$ and for each n, $j = 0, 1, \ldots, 2^{n-1}-1$. These are given by:

$$\begin{aligned} g_{0,0} &= 1; \\ g_{n,j}(s) &= \begin{cases} 2^{(n-1)/2}, & \text{if} \quad s \in [\frac{j}{2^{n-1}}, \frac{j+\frac{1}{2}}{2^{n-1}}\ ), \\ -2^{(n-1)/2}, & \text{if} \quad s \in [\frac{j+\frac{1}{2}}{2^{n-1}}, \frac{j+1}{2^{n-1}}\ ), \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

These functions are known to form a compute ortho-normal system in $L^2[0,1]$.

Define $G_{n,j}(t) = \int_0^t g_{n,j}(s)ds = (\chi_t, g_{n,j})$, where $\chi_t(\cdot)$ is the characteristic function of the interval $[0,t]$, and $(,)$ is the inner product in $L^2[0,1]$. It is helpful to compute the exact expressions for the functions $G_{n,j}(t)$. The maximum value of $G_{n,j}(u)$ occurs at $u = (j+\frac{1}{2})/2^{n-1}$, and $\max\{G_{n,j}(u) \bigm| 0 \le u \le 1\} = 2^{-(n+1)/2}$. Also if $n > 1$, the functions $G_{n,j}$ have disjoint supports.

Now let $\{y_{n,j}\}$ be mutually independent $N(0,1)$ r.v.'s on $(\Omega, \mathcal{A}, P)$, and define the series

$$\sum_{n=0}^{\infty} \sum_{j \in S_n} y_{n,j}(\omega) G_{n,j}(t) \qquad (\omega \in \Omega, t \in T) \ ,$$

where $S_n = \{j \mid 0 \leq j \leq 2^{n-1} - 1\}$ if $n \geq 1$; $S_0 = \{0\}$. The series can be written as $\sum_{n=0}^{\infty} f_n(t,\omega)$ where $f_n(t,\omega) = \sum_{j \in S_n} y_{n,j}(\omega) G_{n,j}(t)$. Let $Y_n(\omega) = \max_{j \in S_n} |y_{n,j}(\omega)|$. Then for any positive number a_n, and $n \geq 1$,

$$P(Y_n > a_n) \leq \sum_{j \in S_n} y_{n,j}\ P(|y_{n,j}| > a_n) \leq C \frac{2^n}{a_n^4} e^{-a_n^2/4}$$

where $C = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} x^4 e^{-x^2/4} dx$, as will be seen after some computation. If we choose $a_n = 2(n \log 2)^{1/2}$, where the logarithm is the natural logarithm, we find

$$\sum_{n=1}^{\infty} P\big(Y_n(\omega) > a_n\big) \leq C \sum_{n=1}^{\infty} \frac{1}{16(\log 2)^2 n^2} < \infty\ .$$

Hence if we define

$$\Omega_0 = \Big\{\omega \ \Big|\ Y_n(\omega) \leq 2(n \log 2)^{1/2} \quad \text{for all sufficiently large} \quad n\Big\}\ ,$$

then by the Borel-Cantelli Lemma, $P(\Omega_0) = 1$. It follows that

$$P\Big\{\omega \ \Big|\ \sum_{n=0}^{\infty} \max_{0 \leq t \leq 1} \ |f_n(t,\omega)| < \infty\Big\} = 1, \quad \text{and} \quad \sum_{n=0}^{\infty} f_n(t,\omega)$$

converges uniformly in t for $\omega \in \Omega_0$ (and thus, a.s.). Finally, define $W_t(\omega)$ to be $\sum_{n=0}^{\infty} f_n(t,\omega)$ when $\omega \in \Omega_0$, and to be equal to 0 $\forall t$ when $\omega \notin \Omega_0$.

It remains to verify that $W = (W_t)$ is a Wiener process on $(\Omega, \mathcal{A}, P)$. If $\omega \in \Omega_0$, $W_t(\omega)$ is a continuous function of t, and thus condition 4 is satisfied. Further the finite dimensional distributions of W are Gaussian. Next

$$E\Big[W_t(\omega) W_s(\omega)\Big] = \lim_{m \to \infty} \left[\Big\{\sum_{n=0}^{m} \sum_{j \in S_n} y_{n,j}(\omega) G_{n,j}(t)\Big\}\Big\{\sum_{n=0}^{m} y_{nj}(\omega) G_{nj}(s)\Big\}\right]$$

$$\sum_{n=0}^{\infty} \sum_{j \in S_n} (\chi_t, g_{nj})(\chi_s, g_{nj}) \ = \ (\chi_t, \chi_s) = \min(t,s),$$

where we used: $E\Big[W_t(\omega) - \sum_{n=0}^{m} \sum_{j\in S_n} y_{nj}(\omega)G_{nj}(t)\Big]^2 \to 0$ as $m \to \infty$. Also $EW_t = 0 \forall t$. Thus the conditions of the second definition of the Wiener process are satisfied.

Next we let $T = R_+$. Define

$$h_{nj}(t) = (2/\pi)^{1/2} \frac{1}{(1+t^2)^{1/2}} g_{nj}\Big(\frac{2}{\pi} \ \arctan t\Big) \text{ for } 0 \le t < \infty \ ,$$

where $j = 0$ if $n = 0$ and $j = 0, 1, \ldots, 2^{n-1} - 1$ if $n \ge 1$. The next step is to verity $\{h_{nj}\}$ is a complete ortho-normal system in $L^2(\mathsf{R}_+)$. We can state this as a lemma.

Lemma 4.1. *$\{h_{nj}\}$ is a complete ortho-normal system in $L^2(\mathsf{R}_+)$.*

Sketch of the Proof. We note that

$$\int_0^\infty h_{nj}(t)h_{mk}(t)dt \ = \ \int_0^1 g_{nj}(t)g_{mk}(t)dt \ = \ \delta_{mn}\delta_{jk} \ .$$

where δ_{ab} is the Kronecker delta, viz. $\delta_{ab} = 1$ if $a = b, \delta_{ab} = 0$ if $a \ne b$. Thus the orthogonal property follows.

The completeness of this system is verified as follows. If $x(t)$ belongs to the class $L_0^1(0,\infty)$ of functions satisfying:

$$\int_0^\infty \frac{|x(t)|}{1+t} dt < \infty \ ,$$

then $\int_0^\infty x(t)^2 dt \ \ge \ \frac{2}{\pi}\Big(\int_0^\infty \frac{|x(t)|}{1+t}dt\Big)^2$, which means that $L^2[0,\infty) \ \subset \ L_0^1[0,\infty)$. Further $x(\cdot) \in L_0^1[0,\infty) \Leftrightarrow \int_0^1 \frac{|x(\tan(\pi s/2))|}{\cos(\pi s/2)}ds < \infty$. Now let $x(\cdot) \in L_0^1[0,\infty) \ni \int_0^\infty x(t)h_{nj}(t)dt = 0$ for $n = 0, j = 0$, and $j = 0, 1, \ldots, 2^{n-1} - 1$ for $n \ge 1$. A little calculation shows that

$$\int_0^\infty x(t)h_{nj}(t)dt = \Big(\frac{\pi}{2}\Big)^{1/2} \int_0^1 \frac{|x(\tan(\pi s/2))|}{\cos(\pi s/2)} g_{nj}(s)ds = 0 \ .$$

The system $\{g_{nj}\}$ is complete in $L^1[0,1]$ i.e. if $x \in L^1[0,1]$ and $\int_0^1 x(t)g_{nj}(t)dt = 0 \ \forall g_{nj}$ then $x = 0$ a.e. This implies that $\{h_{nj}\}$ is complete in $L_0^1[0,\infty)$, hence in $L^2[0,\infty)$. This proves the Lemma.

Now let $H_{nj}(t) = \int_0^t h_{nj}(u)du$. The next lemma gives a convenient estimate concerning the functions $H_{nj}(t)$.

Lemma 4.2. $\sum_{j=0}^{2^{n-1}-1} |H_{nj}(t)| \le \pi 2^{-n/2}(a+\frac{1}{2})$ *for* $t \in [0,a]$, $a < \infty$.

Proof. Let $\delta = \left(\frac{2}{\pi}\right) \arctan t$. Then

$$\begin{aligned} H_{nj}(t) &= \int_0^t h_{nj}(u)du = \sqrt{\frac{\pi}{2}} \int_0^\delta \frac{g_{nj}(s)ds}{\cos(\pi s/2)} ds \\ &= \sqrt{\frac{\pi}{2}} \frac{G_{nj}(\delta)}{\cos(\pi s/2)} - \left(\frac{\pi}{2}\right)^{3/2} \int_0^\delta G_{nj}(s) \frac{\sin(\pi s/2)}{\cos^2(\pi s/2)} ds \ . \end{aligned}$$

Now since $\sum_{j=0}^{2^{n-1}-1} G_{nj}(s) \le 2^{-(n+1)/2}$, we find

$$\begin{aligned} \sum_{j=0}^{2^{n-1}-1} |H_{nj}(t)| &\le \sqrt{\frac{\pi}{2}} 2^{-(n+1)/2} [2(t^2+1)^{1/2} - 1] \\ &< \sqrt{\pi} \cdot 2^{-n/2}(a + \frac{1}{2}) \ . \end{aligned}$$

This proves the Lemma.

Now consider the series $\sum_{n=0}^{\infty} \sum_{j \in S_n} y_{nj}(\omega) H_{nj}(t)$, which we write as $\sum_{n=0}^{\infty} f_n(t,\omega)$ with $f_n(t,\omega) = \sum_{j \in S_n} y_{nj}(\omega) H_{nj}(t)$. Set $K_a = [0,a]$; then

$$\begin{aligned} \max_{t \in K_a} |f_n(t,\omega)| &\le Y_{nj}(\omega) \sum_{j \in S_n} y_{nj}(\omega) |H_{nj}(t)| \\ &\le Y_{nj}(\omega) \left[\sqrt{\pi} 2^{-n/2}(a + \frac{1}{2})\right] \ . \end{aligned}$$

Earlier it was shown that $P(\Omega_0) = 1$, and that for $\omega \in \Omega_0$, $\exists n_0(\omega) \ni \forall n \ge n_0(\omega)$

$$\max_{K_a} |f_n(t,\omega)| \le \sqrt{\pi} \cdot (2a+1) 2^{-n/2} \cdot (n \log 2)^{1/2} \ ,$$

hence

$$\sum_{n=0}^{\infty} \max_{K_a} |f_n(t,\omega)| < \infty \quad \text{(a.s.)} ,$$

and the series $\sum\limits_{n=0}^{\infty} \sum\limits_{j \in S_n} y_{nj}(\omega) H_{nj}(t)$ converges uniformly on each interval K_a for $P-$ a.a. ω. Now we define $W_t(\omega)$ to be the sum of $\sum\limits_{n=0}^{\infty} \sum\limits_{j \in S_n} y_{nj}(\omega) H_{nj}(t)$ if $\omega \in \Omega_0$, and equal to $0 \forall\, t \in \mathbb{R}_+$ if $\omega \notin \Omega_0$. Then $W = (W_t)$, $t \in \mathbb{R}_+$, is a Wiener process on $(\Omega, \mathcal{A}, P)$. This proves the Lemma.

§5. Jump processes - a theorem of Skorokhod

A. Heuristic motivation.

In this section we shall consider convergence to a process having at the worst simple jump discontinuities. Such a process has been latterly described as CADLAG - i.e. continuous from the right ("continue à droite") and having a left limit ("limites à gauche") at each point. Our objective is to present some work of A.V. Skorokhod (cf. [48], [56], [57]). For basic information concerning Markov chains and Markov processes cf. Rosenblatt [50], [51].

The motivation is as follows. Consider a sequence of Markov chains: $\xi_0^{(n)}, \xi_1^{(n)}, \ldots, \xi_n^{(n)}$, and a sequence of partitions:

$$0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{n+1}^{(n)} = 1$$

of the interval $[0,1]$. With each of these Markov chains we associate a random process $\xi^{(n)}(t) = \xi_k^{(n)}$ for $t \in [t_k^{(n)}, t_{k+1}^{(n)}]$, which is a step function with stochastic discontinuities at the points $t_k^{(n)}$. If $\xi_{k+1}^{(n)} - \xi_k^{(n)} \to 0$ uniformly in probability w.r.t. k as $n \to \infty$ and $\max_k (t_{k+1}^{(n)} - t_k^{(n)}) \to 0$, then the process $\xi^{(n)}(t)$ can be considered to be approximately stochastically continuous. We would like to know specific conditions under which $\xi^{(n)}(t)$ will converge, in some appropriately defined sense, as $n \to \infty$, to a process $\xi(t)$ which is a solution of a stochastic equation (cf. [25]), so that under these conditions, $\xi^{(n)}(t)$ can be considered an approximate solution of this stochastic equation.

The processes $\xi^{(n)}(t)$ and $\xi(t)$, as solutions of stochastic equations, with probability one do not have discontinuities of the second kind: the $\xi^{(n)}(t)$ are continuous from the

right, and the $\xi(t)$ can be considered to have the same property. We shall consider some convergence results for such processes.

B. The space $D[0,1]$; Skorokhod's J-topology.

We shall define the basic space of functions and a convenient topology on it. Denote by $D[0,1]$ the space of real-valued functions $x(t)$ defined on the interval $I = [0,1]$, and having the following properties.

1. at each $t \in I$, the left-hand limits $x(t-0)$ exists, and $x(1-0) = x(1)$,
2. at each $t \in [0,1), x(t)$ is right-continuous.

Skorokhod (cf. [57]) introduced and investigated several topologies on this space $D[0,1]$, but we shall use only one of these, denoted in [57] by J_1. For convenience we shall abbreviate this symbol to J. Before defining this topology, it is convenient to note some simple properties of the functions $x(t) \in D[0,1]$. These will be stated as lemmas. If a function $x(t) \in D[0,1]$ has a discontinuity at say $t_0 \in I$, then the magnitude of the jump at this point is clearly $|x(t_0-0) - x(t_0+0)|$; for convenience we shall call this magnitude the "discontinuity" at t_0.

Lemma 5.1. *If $x(t) \in D[0,1]$ then for any $\varepsilon > 0$, $\exists$ only a finite number of values of t such that the discontinuity at t is $> \varepsilon$.*

Proof. For suppose $\exists$ infinite sequence $\{t_k\}$ with $t_k \to t_0$, say, and $|x(t_k+0) - x(t_k - 0)| > \varepsilon$, then at t_0, $x(t)$ would not have a limit either from the right or from the left.

Lemma 5.2. *Suppose $t_1, t_2, \ldots, t_k$ are all the points in I where $x(t) \in D[0,1]$ has discontinuities $\geq \varepsilon$, for some $\varepsilon > 0$. Then $\exists \delta > 0 \ni$ if $|t' - t''| < \delta$ and if both t', t'' belong to the same one of the subintervals $(0,t_1)$, (t_1,t_2), $\ldots$, $(t_k,1)$, then $|x(t') - x(t'')| < \varepsilon$.*

Proof. For suppose $\exists$ sequences $\{t'_n\}$, $\{t''_n\}$ both converging so some point t_0 and belonging to the same one of the intervals $(0, t_1)$, ..., $(t_k, 1)$ and $\ni$ $|x(t'_n) - x(t''_n)| \geq \varepsilon$. Then t'_n, t''_n must lie on opposite sides of t_0 (otherwise $|x(t'_n) - x(t''_n)| \geq \varepsilon$ would be impossible!), hence $|x(t_0 + 0) - x(t_0 - 0)| \geq \varepsilon$. Then t_0 must be one of $t_1, t_2, \ldots, t_k$; but this would contradict the conclusion above that t'_n, t''_n belong to the same one of the intervals $(0, t_1)$, ..., $(t_k, 1)$.

Lemma 5.3. *If $x(t) \in D[0,1]$, then $\forall \eta > 0$ $\exists \delta > 0$ $\ni$ every $t \in I$ satisfies one of the inequalities:*

$$\sup_{0<t_1-t<\delta} |x(t) - x(t_1)| < \eta ,$$

$$\sup_{0<t-t_1<\delta} |x(t) - x(t_1)| < \eta .$$

This can be stated a little more conveniently as

Lemma 5.3′. *If $x(t) \in D[0,1]$, then*

$$\sup_{-\delta<t_1-t<0<t_2-t<\delta} \min\Big[|x(t_1) - x(t)|;\ |x(t) - x(t_2)|\Big] \to 0 \quad \text{as} \quad \delta \to 0 . \qquad (*)$$

The justification of either statement is in Lemma 2.

The next lemma is almost the converse of the assertion of Lemma 5.3.

Lemma 5.4. *If a function $x(t) \in D[0,1]$ satisfies the statement of Lemma 5.3′ then $\exists \bar{x}(t) \in D[0,1]$ which coincides with $x(t)$ at all its points of continuity.*

For suppose $x(t)$ satisfies the statement of Lemma 5.3′. Then $x(t)$ must have, at each $t \in I$, a limit from the left and a limit from the right. Otherwise $\exists \eta > 0$ and three points $t_1 < t_2 < t_3$ in $I \ni |x(t_1) - x(t_2)| > \eta$ and $|x(t_2) - x(t_3)| > \eta$. But this contradicts

(*). Hence $x(t)$ must equal either $x(t-0)$ or $x(t+0)$. Setting $\bar{x}(1) = \lim_{t \to 1_-} x(t)$ and $\bar{x}(t) = \lim_{\substack{t' \to t \\ t' > t}} x(t')$.

Skorokhod's space $D[0,1]$ and his J-topology on this space are meant to be a generalisation of the space $C[0,1]$ of continuous functions on I with the uniform topology. The following definition is meaningful for functions $x(t) \in C[0,1]$.

Definition. The sequence of functions $\{x_n(t)\}_{n=1}^{\infty}$ *converges uniformly to* $x(t)$ *at the point* $t_0 \in I$ if $\forall \varepsilon > 0 \exists \delta > 0 \ni$

$$\overline{\lim_{n \to \infty}} \sup_{|t-t_0|<\delta} |x_n(t) - x(t)| < \varepsilon \ .$$

Facts 5.5.

(1) If $\{x_n t\}$ converges uniformly to $x(t)$ at every point of some closed set then $\{x_n t\}$ converges uniformly to $x(t)$ on this whole set.

(2) A necessary condition for convergence of $\{x_n(t)\}$ to $x(t)$ in the J-topology (defined further on) is: the sequence $\{x_n(t)\}$ converges to $x(t)$ uniformly at every point of continuity of $x(t)$.

The J-topology (defined below) reduces to uniform convergence for continuous functions. The difference appears when we consider the behaviour of $x_n(t)$ in the neighbourhood of points of discontinuity of $x(t)$. Uniform convergence would imply the existence of a number N such that for all $n \geq N$ the points of discontinuity of $x_n(t)$ would coincide with the points of discontinuity of $x(t)$. This would mean that if t were considered to be time, then we would have to assume the existence of an instrument capable of measuring time exactly; physically this is impossible. It seems a little more natural to

suppose that functions which we can obtain by small deformations of the time scale are close to one another. Thus one is naturally led to consider Skorokhod's J-topology.

Skorokhod's concept of convergence appropriate for jump-processes as also for continuous processes, is as follows.

Definition 5.6. A sequence $\{x_n(t)\} \in D[0,1]$ is J-*convergent* to $x_0(t) \in D[0,1]$ if $\exists$ sequence $\{\lambda_n(t)\}$ of continuous monotonically increasing functions $\ni \lambda_n(0) = 0$, $\lambda_n(1) = 1$ and satisfying:

$$\lim_{n\to\infty} \sup_t \Big(\big|x_n\big(\lambda_n(t)\big) - x_0(t)\big| + \big|\lambda_n(t) - t\big|\Big) = 0 .$$

This type convergence will be denoted by:

$$x_n(t) \xrightarrow{\mathrm{J}} x_0(t), \quad \text{or} \quad x_0(t) = \mathrm{J}-\lim_{n\to\infty} x_n(t) .$$

An alternate equivalent definition is stated below; however a convenient notation should be first stated. For a given $\varepsilon > 0$ write

$$x^\varepsilon(t) = x(t) - \sum_{s\le t} [x(s) - x(s-0)]$$

where the understanding is that the sum is over those discontinuities for which $x(s) - x(s-0) > \varepsilon$.

Definition 5.6$'$. A sequence of functions $\{x_n(t)\}$ in $D[0,1]$ is said to be J-*convergent to* $x_0(t) \in D[0,1]$ if

(a) for every $\varepsilon > 0 \ni x_0(t)$ does not have jumps in absolute value equal to $\varepsilon, x_n(t) - x_n^\varepsilon(t) \to x_0(t) - x_0^\varepsilon(t)$ for almost all $t \in I$;

(b) $\lim\limits_{\varepsilon\to 0} \overline{\lim\limits_{n\to\infty}} \sup\limits_{t\in I} |x_n^\varepsilon(t) - x_0^\varepsilon(t)| = 0.$

Now suppose $x(t) \in D[0,1]$. Then for every $C > 0$, write

$$\Delta_C\big(x(t)\big) = \sup_{\substack{0 \le t' < t'' \le 1 \\ |t'-t''| < C}} \min\big(|x(t') - x(t)|; \quad |x(t) - x(t'')|\big)$$
$$+ \sup_{0 \le h \le C} \big(|x(h) - x(0)| + |x(1) - x(1-h)|\big) .$$

Then each function $x(t) \in D[0,1]$ satisfies:

$$\lim_{C \to 0} \Delta_C\big(x(t)\big) = 0 .$$

Now define $\rho_C\big(x(t), y(t)\big)$ $\forall x(t), y(t) \in D[0,1]$, and $\forall C \in 0 < C < 1$, by:

$$\rho_C\big(x(t), y(t)\big) = \Delta_C\big(x(t)\big) + \Delta_C\big(y(t)\big)$$
$$+ \sup_{0 \le t \le 1 - \frac{C}{2}} \inf_{s \in [t, t+\frac{C}{2}]} |x(s) - y(s)| .$$

This quantity $\rho_C\big(x_n(t), x_0(t)\big)$ is used in the following convergence theorem.

C. Skorokhod's theorem

Theorem 5.7.

(i) If J-$\lim_{n\to\infty} x_n(t) = x_0(t)$, *then* $\lim_{C\to 0} \overline{\lim_{n\to\infty}} \rho_C\big(x_n(t), x_0(t)\big) = 0$.

(ii) If $\exists$ *sequence* $\{C_n\}$ *with* $C_n \to 0 \ni \lim_{n\to\infty} \rho_{C_n}\big(x_n(t), x_0(t)\big) = 0$, *then* $x_0(t) =$ J-$\lim_{n\to\infty} x_n(t)$.

Proof.

(i) Suppose $x_0(t) =$ J-$\lim_{n\to\infty} x_n(t)$. Then from the second definition of J-convergence it follows that $x_n(t) \to x_0(t)$ at every point of continuity of $x_0(t)$. Hence, if $C \in (0,1)$:

$$\sup_{0 \le t \le 1-C} \inf_{s \in [t, t+C]} |x_n(s) - x_0(s)| \to 0 \qquad \text{as} \quad n \to \infty .$$

Hence $\Delta_C\big(x_0(t)\big) \to 0$ as $C \to 0$. Thus we only have to show that:

$$\lim_{C\to 0} \varlimsup_{n\to\infty} \Delta_C\big(x_n(t)\big) = 0 .$$

Suppose the $\lambda_n(t)$ satisfy the conditions of Definition 2. Suppose $\sup_t |\lambda_n(t) - t| < \frac{C}{2}$, then

$$\Delta_C\big(x_n(t)\big) \le \Delta_{\frac{C}{2}}\big(\big(x_0(t)\big) + 3 \sup_t |x_n\big(\lambda_n(t)\big) - x_0(t)|$$

and hence

$$\varlimsup_{n\to\infty} \Delta_C\big(x_n(t)\big) \le \Delta_{\frac{C}{2}}\big(\big(x_0(t)\big) ,$$

and therefore

$$\lim_{C\to 0} \varlimsup_{n\to\infty} \Delta_C\big(x_n(t)\big) \le \lim_{C\to 0} \Delta_{\frac{C}{2}}\big(\big(x_0(t)\big) = 0 .$$

Thus part (i) of the theorem follows.

To prove part (ii) of the theorem we need the following two lemmas.

Lemma 5.8. *Let $\varepsilon > 0$, $x(t), y(t) \in D[0,1]$, and $x(t_1) - x(t_1 - 0)| > \varepsilon$. Suppose further, for some $h > 0$, $\rho_h\big(x(t), y(t)\big) < \mu$ where $\mu < \frac{\varepsilon}{7}$. Then $\exists t' \in (t_1 - \frac{h}{2}, t_1 + \frac{h}{2}) \ni$*

$$t \in (t' - h, t') \Rightarrow |y(t' - 0) - x(t_1 - 0)| < 3\mu ,$$

$$|y(t') - x(t_1)| < 3\mu ,$$

$$\text{and } |y(t) - y(t' - 0)| \le \mu ;$$

and

$$t \in [t', t' + h] \Rightarrow |y(t) - y(t')| < \mu .$$

Proof of Lemma 5.8. $\exists s_1 \in (t_1 - \frac{h}{2}, t_1)$ and $\exists s_2 \in (t_1, t_1 + \frac{h}{2})$, $\ni |y(s_1) - x(s_1)| < \mu$ and $\ni |y(s_2) - x(s_2)| < \mu$. Now $\Delta_h\big(x(t)\big) < \mu$, hence

$$\min \{|x(s_1) - x(t_1 - 0)|; \qquad |x(t_1) - x(t_1 - 0)|\} < \mu ;$$

and

$$\min\{|x(s_2)-x(t_1)|; \qquad |x(t_1)-x(t_1-0)|\} < \mu \ .$$

It follows that $|x(s_1)-x(t_1-0)| < \mu$ and $|x(s_2)-x(t_1)| < \mu$. Hence $|y(s_1)-x(t_1-0)| < 2\mu$, and $|y(s_2)-x(t_1)| < 2\mu$, and thus $|y(s_1)-y(s_1)| > \varepsilon - 4\mu > 3\mu$. Further

$$\min_{s_1 \le s' < t < s'' \le s_2} \big(|y(s')-y(t)|; |y(t)-y(s'')|\big) < \mu \ ,$$

hence $\exists t' \in [s_1, s_2] \ni$

$$s_1 \le t < t' \ \Rightarrow \ |y(s_1)-y(t)| < \mu$$

and

$$t' \le t \le s_2 \ \Rightarrow \ |y(s_2)-y(t)| < \mu \ .$$

Therefore $|y(s_1)-y(t'-0)| \le \mu$ and $|y(s_2)-y(t')| < \mu$; also $|y(t'-0)-x(t_1-0)| < 3\mu$ and $|y(t')-x(t_1)| < 3\mu$

Now $|y(t'-0)-y(t')| > \varepsilon - 6\mu > \mu$ and

$$\min_{t'-h<t<t'} \big(|y(t)-y(t'-0)|; |y(t'-0)-y(t')|\big) < \mu,$$

hence $\big(|y(t)-y(t'-0)| \le \mu$. Similarly $|y(t)-y(t')| < \mu$ for $t \in [t', t'+h]$. The Lemma follows.

Lemma 5.9. *Suppose the jumps of $x(t)$ and $y(t)$ do not exceed ε, where $\varepsilon > 0$ is given, and suppose for some $h < \frac{1}{2}$, $\rho_h\big(x(t), y(t)\big) < \mu$, then $\sup_t |x(t)-y(t)| < 2\varepsilon + 5\mu$.*

Proof of Lemma 5.9. We shall first show that

$$|t_1 - t_2| < h \ \Rightarrow \ |x(t_1)-x(t_2)| < \varepsilon + 2\mu \ .$$

Let $t_1 < t_2$ and $t' \in (t_1, t_2) \ni$ if $s < t'$ and $|x(t_1) - x(t')| > \mu$ then $|x(t_1) - x(s)| \le \mu$. Then

$$|x(t_1) - x(t')| \le |x(t_1) - x(t'-0)| + |x(t'-0) - x(t')| \le \varepsilon + \mu \ .$$

However $\min\big(|x(t_1) - x(t')|;\ |x(t') - x(t_2)|\big) < \mu$, hence $|x(t') - x(t_2)| < \mu$, and therefore $|x(t_1) - x(t_2)| < \varepsilon + 2\mu$.

Similarly $|t_1 - t_2| < h \Rightarrow |y(t_1) - y(t_2)| < \varepsilon + 2\mu$. If $t \in [0,1]$ then $\exists\ t' \ni |t - t'| < \frac{h}{2}$ and $|y(t') - x(t')| < \mu$. Therefore

$$|y(t) - x(t)| \le |y(t') - x(t')| + |y(t') - y(t)| + |x(t') - x(t)| < 2\varepsilon + 5\mu \ .$$

This proves the lemma.

We shall now return to the proof of Theorem 5.7 (ii).

Proof of (ii) of Theorem 5.7. Suppose $\rho_{C_n}\big(x_n(t), x_0(t)\big) \to 0$ and $C_n \to 0$. Choose $\varepsilon > 0 \ni x_0(t)$ does not have jumps equal in absolute value to ε. Then $\exists \mu \ni \mu < \frac{\varepsilon}{7}$ and $x_0(t)$ does not have jumps whose absolute values fall in the interval $[\varepsilon - 6\mu, \varepsilon + 6\mu]$. Suppose $t_1 < t_2 < \cdots < t_k$ are all the points at which the jumps of $x_0(t)$ exceed ε in absolute value. Let $\delta = \min\limits_{0 \le i \le k} (t_{i+1} - t_i)$, where $t_{k+1} = 1$. Let the positive integer n be so large that $C_n < \frac{\delta}{2}$ and $\rho_{C_n}\big(x_n(t), x_0(t)\big) < \mu$. Then in the intervals $(t_1 - \frac{C_n}{2}, t + \frac{C_n}{2})$, by Lemma 5.2 $\exists$ points $t_i^{(n)} \ni |x_n(t_i^{(n)} - 0) - x_0(t_i - 0)| < 3\mu$ and $|x_n(t_i^{(n)} - 0) - x_0(t_0)| < 3\mu$. Hence

$$|x_n(t_i^{(n)} - 0) - x_n(t_i^{(n)}| > |x_0(t_i) - x_0(t_i - 0)| - 6\mu > \varepsilon \ .$$

By Lemma 5.6 we conclude that $x_n(t)$ does not have jumps exceeding 2μ in absolute value in the intervals $(t_i^{(n)} - C_n, t_i^{(n)})$, $(t_i^{(n)}, t_i^{(n)} + C_n)$. Hence in each of the intervals

$t_i - \frac{C_n}{2}, t_i + \frac{C_n}{2})$, $\exists$ only one jump of $x_n(t)$ with absolute value exceeding ε. Now $x_n(t)$ cannot have a jump exceeding ε in absolute value at a point $t' \neq t_i^{(n)}$ for any n.

For otherwise by Lemma 5.1 there must be a point $t'' \ni |t' - t''| < \frac{C_n}{2}$ and $|x_0(t'') - x_0(t'' - 0)| > \varepsilon - 6\mu$. But, because of the choice of μ, the inequality $|x_0(t'') - x_0(t'' - 0)| > \varepsilon - 6\mu \Rightarrow |x_0(t'') - x_0(t'' - 0)| > \varepsilon$, hence for some j, $t'' = t_j$. Then the interval $(t_j - \frac{C_n}{2}, t_j + \frac{C_n}{2})$ would have two points t' and $t_j^{(n)}$ at which the jumps of $x_n(t)$ exceed ε in absolute value, which is impossible. Hence

$$x_n(t) - x_n^\varepsilon(t) = \sum_{t_i^{(n)} \leq t} \left(x_n(t_i^{(n)}) - x_n(t_i^{(n)} - 0)\right) ,$$

and

$$x_0(t) - x_0^\varepsilon(t) = \sum_{t_i \leq t} \left(x_0(t_i) - x_0(t_i - 0)\right) .$$

By Lemma 5.8, and the condition that $\rho_{C_n}\left(x_n(t), x_0(t)\right) \to 0$, it is clear that $x_n(t) \to x_0(t)$ at every point of continuity of $x_0(t)$. Hence as $n \to \infty$, $|x_n(t_i^{(n)}) - x_0(t_i)| - |x_n(t_i^{(n)} - 0) - x_0(t_i - 0)| \to 0$. Also because $C_n \to 0$, therefore $t_i^{(n)} \to t_i$. Hence $x_n(t) - x_n^\varepsilon(t) \to x_0(t) - x_0^\varepsilon(t)$ as $t \neq t_i$. Therefore $x_n^\varepsilon(t) \to x_0^\varepsilon(t)$ for all points t of continuity of $x_0^\varepsilon(t)$ (as this is fulfilled for $x_n(t)$). We conclude also that condition (a) of Definition 1 is fulfilled for $x_n(t)$.

We shall now show that when $\Delta_C\left(x(t)\right) < \varepsilon, \Delta_C\left(x^\varepsilon(t)\right) < 2\varepsilon$. If $t_1 < t_2 < t_3$, with $t_3 - t_1 < C$, then

$$\min\left(|x^\varepsilon(t_1) - x^\varepsilon(t_2)|, |x^\varepsilon(t_2) - x^\varepsilon(t_3)|\right) \leq \Delta_C\left(x(t)\right)$$

provided $x(t)$ has no jump exceeding ε in absolute value in (t_1, t_3); and if a jump with

absolute value exceeding ε does exist at, say, $t' \in (t_1, t_2)$, then

$$\begin{aligned}
\min\big(|x^\varepsilon(t_1) - x^\varepsilon(t_2)|, |x^\varepsilon(t_2) - x^\varepsilon(t_3)|\big) &\leq |x^\varepsilon(t_2) - x^\varepsilon(t_3)| \\
&\leq |x(t_2) - x(t')| + |x(t_3) - x(t')| \\
&= \min\big(|x(t'-0) - x(t')|, |(x(t') - x(t_2)|\big) \\
&\quad + \min |x(t'-0) - x(t')|, |(x(t') - x(t_3)| \\
&\leq 2\Delta_C\big(x(t)\big) \ ;
\end{aligned}$$

and a similar inequality holds if $x(t)$ has a jump exceeding ε in absolute value in (t_2, t_3). Therefore $\lim_{n\to\infty} \Delta_{C_n}\big(x_n^\varepsilon(t)\big) = 0$.

Now consider $\inf_{\bar{t}<t<\bar{t}+\frac{C_n}{2}} |x_n^\varepsilon(t) - x_0^\varepsilon(t)|$. If there is no point $t_i, t_i^{(n)}$ in $\bar{t}, \bar{t} + \frac{C_n}{2})$, then

$$\begin{aligned}
\inf_{\bar{t}\leq t<\bar{t}+\frac{C_n}{2}} |x_n^\varepsilon(t) - x_0^\varepsilon(t)| < & \inf_{\bar{t}\leq t<\bar{t}+\frac{C_n}{2}} |x_n(t) - x_0(t)| \\
& \sum_{i=1}^{k} \big(|x_n(t_i^{(n)} - x_0(t_i)| + |x_n(t_i^{(n)} - 0) - x_0(t_i - 0)|\big) \\
& (6k+1)\rho_{C_n}\big(x_n(t), x_0(t)\big) \ .
\end{aligned}$$

If $\exists t_i$ with $t_i^{(n)} < t_i$ in $(\bar{t}, \bar{t} + \frac{C_n}{2})$ then since $|t_i^{(n)} - t_i| < \frac{C_n}{2}$ for $t' \in (t_i, \bar{t} + \frac{C_n}{2})$, we have $|t' - t_i| < \frac{C_n}{2}$, $|t_i^{(n)} - t'| < C_n$, and thus

$$\begin{aligned}
|x_n^\varepsilon(t') - x_0^\varepsilon(t')| &\leq |x_n(t') - x_0(t')| + 6k\rho_{C_n}\big(x_n(t), x_0(t)\big) \\
&\leq |x_n(t') - x_n(t_i^{(n)})| + |x_0(t') - x_0(t_i)| \\
&\quad + \ |x_n(t_i^{(n)}) - x_0(t_i)| + 6k\rho_{C_n}\big(x_n(t), x_0(t)\big) \\
&< (6k+5)\rho_{C_n}\big(x_n(t), x_0(t)\big) \ .
\end{aligned}$$

Using the symmetry of $t_i^{(n)}$ and t_i, we conclude that

$$\inf_{\bar{t}\leq t<\bar{t}+\frac{C_n}{2}} |x_n^\varepsilon(t) - x_0(t)| \leq (6k+5)\rho_{C_n}\big(x_n(t), x_0(t)\big) \ .$$

Hence, if $C_n < \frac{\delta}{2}$, and $\rho_{C_n}\big(x_n(t), x_0(t)\big) < \mu$, then

$$\rho_{C_n}\big(x_n^\varepsilon(t), x_0^\varepsilon(t)\big) \leq (6k+7)\rho_{C_n}\big(x_n(t), x_0(t)\big) \ .$$

By Lemma 5.9

$$\sup_t \ |x_n^\varepsilon(t) - x_0^\varepsilon(t)| \leq 2\varepsilon + 5(6k+7)\rho_{C_n}\big(x_n(t), x_0(t)\big) \ ,$$

and condition (b) of Definition 5.6′ is fulfilled. Hence

$$x_0(t) \ = \ \mathsf{J}\text{–} \lim_{n\to\infty} \ x_n(t) \ .$$

This completes the proof of Skorokhod's theorem.

APPENDIX 1

Topological vector spaces

This appendix aims at presenting the basic ideas and results concerning functional analysis which are needed for this monograph. For proofs we refer the reader to e.g., Rudin [52], or Köthe [29].

The letters $\mathbf{R}$ and $\mathbf{C}$ shall denote the field of real numbers and the field of complex numbers respectively. We shall use the letter $\mathbf{K}$ to denote either $\mathbf{R}$ or $\mathbf{C}$, and an element of $\mathbf{K}$ is called a *scalar*. A vector space over $\mathbf{K}$ is a set X consisting of elements called *vectors*, and in which two operations, viz. *addition* and *scalar multiplication* are defined, with the usual algebraic properties:

(a) for every pair of vectors $x, y \in X, \exists$ a vector $x + y \in X \ni$

$$x + y = y + x, \qquad \text{and} \qquad x + (y + z) = (x + y) + z \ ;$$

X contains a unique vector 0 (called the *zero vector* of X) $\ni x + 0 = x \forall x \in X$, and to each $x \in X \exists |$ vector $-x \ni x + (-x) = 0$;

(b) for every pair (α, x), $\alpha \in \mathbf{K}$, $x \in X, \exists$ vector $\alpha x \in X$ satisfying:

$$1 \cdot x = x; \qquad \alpha(\beta x) = (\alpha\beta)x;$$

$$\alpha(x + y) = \alpha x + \alpha y; \qquad (\alpha + \beta)x = \alpha x + \beta x \ .$$

Note: the symbol 0 also denotes the zero element in $\mathbf{K}$.

X will be called *real* if $\mathbf{K} = \mathbf{R}$, or *complex* if $\mathbf{K} = \mathbf{C}$. Suppose $A \subset X$, $B \subset X$,

$x \in X$, and $\lambda \in \mathbf{K}$. Then

$$x \pm A \quad \text{means the set} \quad \{x \pm a \mid a \in A\};$$

$$A + B \quad \text{means the set} \quad \{a + b \mid a \in A, b \in B\};$$

$$\lambda A \quad \text{means the set} \quad \{\lambda a \mid a \in A\}.$$

Thus $-A$ denotes the set $\{-a \mid a \in A\}$. **Note:** $2A$ may be $\neq A + A$.

A *subspace* of X is a set $Y \subset X \ni Y$ is a vector space with the same operations. $Y \subset X$ is a subspace if and only if $0 \in Y$ and

$$\alpha Y + \beta Y \subset Y \qquad \forall \alpha, \beta \in \mathbf{K} .$$

A *convex set* in X is a subset $C \subset X \ni$

$$tC + (1-t)C \subset C \quad \forall t \in [0,1] .$$

A set $B \subset X$ is *balanced* if $\alpha B \subset B \; \forall \alpha \in \mathbf{K} \ni |\alpha| \leq 1$. A vector space X is n-*dimensional* (or has *dimension* n) if $\exists$ basis $\{u_1, \ldots, u_n\}$ in X, i.e. if and only if each $x \in X$ can be uniquely expressed as

$$x = \alpha_1 u_1 + \cdots + \alpha_n u_n, \alpha_1, \ldots, \alpha_n \in \mathbf{K} .$$

A *finite dimensional* X is one which is n-dimensional for some integer $n > 0$.

Before stating the definition of a topological vector space it appears only appropriate to state briefly some standard vocabulary concerning topological spaces. A *topological space* is a set X in which a special collection τ of subsets (called *open* sets) is specified, with the following properties:

(i) the entire set X is open

(ii) the empty set ϕ is open,

(iii) the intersection of any two open sets is open, and

(iv) the union of an arbitrary collection of open sets is open.

A special family τ of sets with these properties is called a *topology* on X. At times for the sake of precision, the topological space which the family τ turns X into, is denoted by (X, τ).

In a topological space (X, τ), a set E is *closed* if its complement is open. The *closure* of a set E, denoted by $\bar{E}$, is the intersection of all closed sets containing E. The *interior* of a set E, denoted by E°, is the union of all the open sets contained in E. A *neighbourhood* of a point $x \in X$ is any open set containing x. (X, τ) is called a *Hausdorff space* and τ is called a *Hausdorff topology on* X if distinct points of X have disjoint neighbourhoods. A set $K \subset X$ is *compact* if every open covering of K has a finite subcovering. A subfamily τ' of τ is called a *base for* τ if every member of τ is a union of members of τ', or equivalently if for any set $U \in \tau$, and for any point $x \in U$, $\exists$ set $V \in \tau'$ $\ni x \in V \subset U$. A family γ of neighbourhoods of a point $x \in X$ is called a *local base* at x if every neighbourhood of x contains a member of γ. If a topology τ is induced by a metric d, we say that d and τ are *compatible* with each other.

A *topological vector space* (T.V.S.) is a vector space X with a topology τ on $X \ni$

(a) every single-point set in X is a closed set;

(b) the vector space operations are continuous with respect to τ.

A subset $E \subset X$ is *bounded* if for each neighbourhood V of 0 in X $\exists s > 0 \ni \forall t > s$,

$E \subset tV$.

Translation and multiplication operators. Suppose X is a TVS. To each $a \in X$ we associate the *translation* operator T_a defined by: $T_a x = a + x, x \in X$; and to each scalar λ we associate the *multiplication* operator $M_\lambda : M_\lambda x = \lambda x, x \in X$. Then these two operators T_a and M_λ both homeomorphisms of X onto X.

This last statement implies the following: every vector topology on X is *translation invariant* i.e. a set $E \subset X$ is open if and only if for each $a \in X$, $a + E$ is open.

In a T.V.S. X the term *local base* means a local base of neighbourhoods at 0. Thus a local base of a T.V.S. X is a collection $\mathcal{B}$ of neighbourhoods of 0 such that every neighbourhood of 0 contains a member of $\mathcal{B}$. The open sets of X are then precisely those that are unions of translates of members of $\mathcal{B}$.

A metric d on a vector space X is *translation invariant* if $d(x + z, y + z) = d(x, y)$ $\forall x, y, z \in X$.

The following definition explains some of the types of T.V.S.'s that we might encounter. X here denotes a T.V.S. with topology τ.

Definition.

(a) X is *locally convex* if $\exists$ local base $\mathcal{B}$ consisting of convex subsets.

(b) X is *locally bounded* if 0 has a bounded neighbourhood.

(c) X is *locally compact* if 0 has a neighbourhood with compact closure.

(d) X is *metrisable* if τ is induced by a metric d.

(e) X is an *F-space* if τ is induced by a complete invariant metric.

(f) X is a *Fréchet space* if X is a locally convex *F-space*.

(g) A *norm* on a vector space X is a non negative valued function denoted by $\|x\|$, having the properties:

$$\|x\| = 0 \qquad \text{only if} \quad x = 0$$

$$\|\alpha x\| = |\alpha|\|x\| \quad \text{if} \quad \alpha \in \mathbf{K}; \quad x \in X \ ,$$

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in X \ .$$

A vector space X with a norm on X is called a *normed linear space*. If a vector space X is normed then $d(x, y) = \|x - y\|$ defines a distance (or metric) on X. If X is complete w.r.t. this metric, X is called a *Banach space.*

(h) A T.V.S. X is *normable* if $\exists$ norm on $X \ni$ the metric induced by the norm on X is compatible with the topology on X.

(k) A T.V.S. X has the *Heine-Borel property* if every closed and bounded subset of X is compact.

Theorem. *If $\mathcal{B}$ is a local base for a T.V.S. X then every member of $\mathcal{B}$ contains the closure of some member of $\mathcal{B}$.*

Hence:

Corollary. *Every T.V.S. is a Hausdorff space.*

Theorem. *In a T.V.S. X*

(a) every neighbourhood of 0 contains a balanced neighbourhood of 0.

(b) Every convex neighbourhood of 0 contains a convex balanced neighbourhood of 0.

Thus

Theorem.

(a) Every T.V.S. has a balanced local base.

(b) *Every locally convex space has a balanced convex local base.*

Suppose X and Y are vector spaces over the same field $\mathbf{K}$. A mapping $T : X \to Y$ is called *linear* if

$$T(\alpha x + \beta y) \;=\; \alpha Tx + \beta Ty$$

$\forall x, y \in X$ and $\forall \alpha, \beta \in \mathbf{K}$. For a linear mapping we often write Tx instead of $T(x)$. A linear mapping $T : X \to \mathbf{K}$ is called a *linear functional.*

Theorem. *Let X and Y be T.V.S.s. If $T : X \to Y$ is continuous at 0 then T is continuous, and in fact uniformly continuous, i.e., for each neighbourhood W of 0 in Y, $\exists$ neighbourhood V of 0 in $X \ni$*

$$y - x \in V \Rightarrow Ty - Tx \in W \ .$$

For a linear functional on a T.V.S., the following is true.

Theorem. *Suppose F is a linear functional on a T.V.S. X, $\ni Tx \neq 0$ for some $x \in X$. Then the following four statements are equivalent:*

(a) *F is continuous.*

(b) *The null-space $N(F)$ is closed.*

(c) *$N(F)$ is not dense in X.*

(d) *F is bounded in some neighbourhood of 0.*

The simplest models of Banach spaces are the standard real of complex n-dimensional Euclidean spaces $\mathbf{R}^n$ or $\mathbf{C}^n$ over $\mathbf{R}^1$ or $\mathbf{C}$, respectively, normed by means of the usual Euclidean metric.

For example if $z = (z_1, \dots, z_n)$, $z_i \in \mathbb{C}$ is a point (i.e. vector) in $\mathbb{C}^n$ then $\|z\| = \left(\sum_{j=1}^{n} |z_i|^2\right)^{1/2}$ is a norm on $\mathbb{C}^n$; likewise if $x = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}^1$ is a point (or vector) in $\mathbb{R}^n$, then $\|x\| = \left(\sum_{j=1}^{n} |x_i|^2\right)^{1/2}$ is a norm on $\mathbb{R}^n$.

These are by *no means* the only norms that can be introduced on $\mathbb{R}^n$ or $\mathbb{C}^n$, respectively.

Theorem. *Suppose X is a complex T.V.S., Y is a subspace of X, and $\dim Y = n$ where n is a positive integer. Then*

(a) every isomorphism of Y onto $\mathbb{C}^n$ is a homeomorphism;

(b) Y is closed.

Theorem.

(a) Every locally compact T.V.S. is finite dimensional.

(b) If a T.V.S. X is locally bounded and has the Heine-Borel property then X is finite dimensional.

Before turning on to some of the most useful type of T.V.S.s we shall mention the general characteristics of a bounded linear transformation (or linear mapping). A linear mapping $T : X \to Y$, where X, Y are T.V.S.'s, is *bounded* if T maps bounded sets into bounded.

Theorem. *Suppose X, Y are T.V.S.s and $T : X \to Y$ is a linear mapping. Then among the following four properties of X,*

$$(a) \Rightarrow (b) \Rightarrow (c) ,$$

If further X is metrisable then

$$(c) \Rightarrow (d) \Rightarrow (a) ,$$

so that for a metrisable T.V.S. X, all four statements are equivalent.

(a) *T is continuous;*

(b) *T is bounded;*

(c) *if $x_n \to 0$ then $\{Tx_n, n = 1, 2, 3, \ldots\}$ is bounded;*

(d) *if $x_n \to 0$ then $Tx_n \to 0$.*

Among the most useful kind of T.V.S.s occurring in analysis are the locally convex ones, for the topological structure of a locally convex space X can be specified by a special family of non negative (non linear) functions on X called semi-norms.

A *semi-norm* on a vector space X is a real-valued function $p(\cdot)$ on X with the properties:

(a) $p(x + y) \leq p(x) + p(y)$,

(b) $p(\alpha x) = |\alpha| p(x)$, $\forall x, y \in X$ and $\forall \alpha \in \mathbf{K}$. If further p satisfies

(c) $p(x) \neq 0$ if $x \neq 0$ then p is a norm.

A family $\mathcal{P}$ of semi-norms on X is called *separating* if to each $x \neq 0 \quad \exists$ semi-norm $p \in \mathcal{P} \ni p(x) \neq 0$.

If the vector space X is also an algebra, an algebra semi-norm $p(\cdot)$ on X is a semi-norm which further satisfies

(a) $p(x \cdot y) \leq p(x)p(y) \quad \forall x, y \in X$, and

(b) if X further has a unit e then $p(e)$ is either 1 or 0.

A subset $A \subset X$ is called *absorbing* if each $x \in X$ lies in tA for some $t > 0$. Suppose $A \subset X$ is absorbing; then the *Minkowski functional* $\mu_A(\cdot)$ of A is defined by:

$$\mu_A(x) \;=\; \inf[t > 0 \;\Big|\; x \in tA]\,.$$

We note $\mu_A(x) < \infty$ for every $x \in X$. Also the semi-norms on X will be seen to be precisely the Minkowski functionals of the balanced convex absorbing sets in X.

A semi-norm p on a vector space X has the following properties.

Theorem. *Suppose p is a semi-norm on a vector space X. Then*

(a) $p(0) = 0$.

(b) $|p(x) - p(y)| \leq p(x - y)$.

(c) $p(x) \geq 0$.

(d) $\{x \mid p(x) = 0$ *is a subspace of* $X\}$.

(e) The set $B = \{x \mid p(x) < 1\}$ *is convex, balanced and absorbing, and further* $p = \mu_B$.

Theorem. *Suppose A is a convex absorbing set in a vector space X. Then*

(a) $\mu_A(x + y) \leq \mu_A(x) + \mu_A(y)$.

(b) $\mu_A(tx) = t\mu_A(x)$ *if* $t \geq 0$.

(c) If A is balanced then μ_A is a semi-norm.

(d) If $B = \{x \mid \mu_A(x) < 1\}$, $C = \{x \mid \mu_A(x) \leq 1\}$, *then* $B \subset A \subset C$, *and* $\mu_B = \mu_A = \mu_C$.

The next two theorems clarify the relation between families of semi-norms on a T.V.S. and locally convex topological structures on X. A family $\mathcal{P}$ of semi-norms on X is said to be *separating* if for each $x \neq 0 \exists p \in \mathcal{P} \ni p(x) \neq 0$.

Theorem. *Suppose $\mathcal{B}$ is a convex balanced local base in a T.V.S. X. We associate with every $V \in \mathcal{B}$ its Minkowski functional. Then $\{\mu_V \mid V \in \mathcal{B}\}$ is a separating family of continuous semi-norms on X.*

Theorem. *Suppose $\mathcal{P}$ is a separating family of semi-norms on a vector space X. For*

each $p \in \mathcal{P}$ and each positive integer n define the set

$$V_{p,n} = \{x \mid p(x) < \frac{1}{n}\} .$$

Let $\mathcal{B}$ be the collection of all finite intersections of the sets $V_{p,n}$. Then $\mathcal{B}$ is a convex balanced local base for a topology on X which makes X a locally convex space, in which:

(a) every $p \in \mathcal{P}$ is continuous, and

(b) *a set $E \subset X$ is bounded if and only if every $p \in \mathcal{P}$ is bounded on E.*

Remark.

(a) The real use of the separating property of the family $\mathcal{P}$ of semi-norms in the last theorem is in showing that this property implies that in the T.V.S. X every single-point set is closed. However, we should note that this property of a single-point set being closed is sometimes omitted from the definition of a T.V.S. X.

(b) If $\mathcal{B}$ is a convex balanced local base for the topology τ of a locally convex space X then $\mathcal{B}$ generates a separating family $\mathcal{P}$ of continuous semi-norms on X, and $\mathcal{P}$ in turn induces a topology τ_1 on X. Clearly then $\tau_1 \subset \tau$; as a matter of fact $\tau_1 = \tau$.

(c) Suppose $\mathcal{P}$ is a countable separating family of semi-norms on X, then a complicated theorem explains the construction of a translation invariant metric d on X compatible with the topology on X and such that the open balls defined by d are convex. However a much simpler and direct definition of a compatible translation invariant metric in terms of the countable family $\mathcal{P} = \{p_i\}_{i=1}^{\infty}$ is as follows: let

$$d(x,y) = \sum \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)} ;$$

then d is a metric, invariant and compatible with the topology on X. However the balls which are defined by d need not be convex.

One last theorem which should be stated here is as follows.

Theorem. *A* T.V.S. X *is normable if and only if the origin in* X *has a convex bounded neighbourhood.*

The commonly occurring function spaces provide excellent examples of locally convex spaces.

Example 1. The space $C(\Omega)$. Suppose Ω is a non-empty open set in a Euclidean space $\mathbf{R}^N$. It is known that Ω is the union of a countable family of compact sets $\{K_n\}_{n=1}^{\infty} \ni$ each K_n is non-empty and further each $K_n \subset K_{n+1}^{\circ}$ (i.e. in the interior of K_{n+1}). In fact the sets K_n can be defined as follows. Define

$$W_n = B_n(\infty) \bigcup \{ \bigcup_{a \in \mathcal{C}(\Omega)} B_{1/n}(a)\}$$

where $B_n(\infty) = \{x \in \mathbf{R}^N \mid \|x\| > n\}$, and let $K_n = \mathcal{C}(W_n)$, for $n = 1, 2, \ldots$. Then we can verify that the K_n's form one family of compact sets with the properties stated.

The space $C(\Omega)$ is the vector space of all complex-valued continuous functions on Ω. The family $\mathcal{P} = \{p_n\}_{n=1}^{\infty}$ of semi-norms where

$$p_n(f) = \sup\{|f(x) \mid x \in K_n\}, \quad n = 1, 2, \ldots,$$

satisfies clearly $p_1 \leq p_2 \leq \ldots$. The family $\{V_n\}_{n=1}^{\infty}$ defined by:

$$V_n = \{f \in C(\Omega) \mid p_n(f) < \frac{1}{n}\}, \quad n = 1, 2, \ldots,$$

forms a convex local base for $C(\Omega)$.

Topological Properties of $C(\Omega)$. $C(\Omega)$ is a Fréchet space, in which a set E is bounded if and only if every $p \in \mathcal{P}$ is bounded on E. This space is not locally bounded and hence is not normable.

Example 2. The space $C^\infty(\Omega)$. As in Example 1, let Ω be a non-empty open set in $\mathbf{R}^n$. We shall consider the space $C^\infty(\Omega)$ of complex-valued functions f defined in $\Omega \ni D^\alpha f \in C(\Omega)$ (cf. Example 1) for each multi-index α. Here a multi-index α is an ordered n-tuple: $\alpha = (\alpha_1, \ldots, \alpha_n)$ of non-negative integers $\alpha_i, i = 1, \ldots, n$, and with each multi-index α we associate the differential operator

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$$

of *order* $|\alpha| = \alpha_1 + \cdots + \alpha_n$; if $|\alpha| = 0$ we define $D^\alpha f = f$.

This space $C^\infty(\Omega)$ is customarily endowed with a locally convex topology, as follows. Let $\{K_n\}_{n=1}^\infty$ be a sequence of compact sets in Ω selected as in Example 1, $\ni \Omega = \bigcup_{n=1}^\infty K_n$ and $K_n \subset K_{n+1}^\circ$, $n = 1, 2, \ldots$. The semi-norms p_N on $C^\infty(\Omega)$, $N = 1, 2, \ldots$, are defined by

$$p_N(f) = \max\left\{|D^\alpha f(x)| \;\middle|\; x \in K_N,\ |\alpha| \le N\right\}.$$

By the general theorem mentioned earlier, this countable family of semi-norms defines a locally convex topology on $C^\infty(\Omega)$ which also makes it a Fréchet space with an invariant metric as explained above.

This topology on $C^\infty(\Omega)$ further has the following properties:

(i) for each $x \in \Omega$, the functional $f \to f(x)$, is continuous;

(ii) $C^\infty(\Omega)$ has the Heine-Borel property;

(iii) the space is not locally bounded, hence not normable.

Example 3. Let K be a compact set in $\mathbf{R}^n$ with non-empty interior. The *support* of a complex function f (on any topological space) is defined as the closure of the set

$\{x \mid f(x) \neq 0\}$. D_K is defined as the space of C^∞ functions with support contained in K.

If Ω is an open set containing K, we see that D_K is a subspace of $C^\infty(\Omega)$, and hence has all the properties which $C^\infty(\Omega)$ has (cf. Example 2). To see that D_K satisfies property (iii) (cf. Example 2), we note that the restriction that $K^\circ \neq \phi$ implies that $\dim(D_K) = \infty$ because of the following proposition:

Proposition. Suppose B_1 and B_2 are closed balls in $\mathbf{R}^n \ni B_1 \subset B_2^\circ$; then $\exists \phi \in C^\infty(\mathbf{R}^n) \ni \phi \equiv 1$ on $B_1, \phi \equiv 0$ outside B_2 and generally $0 \leq \phi \leq 1$ everywhere.

We shall end this survey of concepts and results concerning topological vector spaces with some explanation about linear functionals on a T.V.S. X.

The dual space of a T.V.S. X is the space X^* whose elements are the continuous linear functionals on X. The next theorems show that there exist lots of continuous linear functionals on a locally convex space.

Theorem. *Suppose*

(i) M is a subspace of a real vector space X,

(ii) $p : X \to \mathbf{R}$ satisfies: $p(x+y) \leq p(x) + p(y)$, and $p(tx) = tp(x)$ if $x, y \in X$ and $t \geq 0$, and

(iii) $f : M \to \mathbf{R}$ is linear and satisfies: $f(x) \leq p(x) \forall x \in M$. Then $\exists$ linear $F : X \to \mathbf{R} \ni F(x) = f(x) \forall x \in M$ and $-p(-x) \leq F(x) \leq p(x)\ \forall x \in X$.

Theorem. *Suppose M is a subspace of a locally convex T.V.S. X, and $x_0 \in X$. If $x_0 \notin \overline{M}$ then $\exists F \in X^* \ni F(x_0) = 1$ but $F(x) = 0\ \forall x \in M$.*

This Page Intentionally Left Blank

APPENDIX 2

Differential Calculus in Banach spaces

We shall collect here the main ideas and results concerning the differential calculus in Banach spaces which are needed in this monograph. No proofs are provided; for the proofs the reader should refer to Dieudonné: Foundations [9] or S. Lang: Real Analysis [33].

All Banach spaces dealt with are assumed to be real, and mappings are supposed to be from subsets of Banach spaces to Banach spaces which are denoted by letters E, F, G, E_1, E_2, etc. Suppose $U \subset E$ is an open set. A mapping $f : U \to F$ is *(Fréchet) differentiable* at $x \in U$ if $\exists$ continuous linear map $A : E \to F$, and $\exists$ map ψ defined for all $h \in E$ with sufficiently small norm with values in F, satisfying

$$\lim_{h \to 0} \psi(h) = 0$$

and also

$$f(x+h) = f(x) + A(h) + \|h\| \, \psi(h) . \tag{1}$$

We shall assume here that $\psi(0)$ is defined and equals 0. Alternatively we can replace the term $\|h\|\psi(h)$ by a term $\phi(h)$ where $\phi(\cdot)$ is a mapping with the property:

$$\lim_{h \to 0} \frac{\phi(h)}{\|h\|} = 0 . \tag{2}$$

Clearly if f is differentiable at x then it is continuous at x. Also if the continuous linear mapping A exists satisfying (1) then it is uniquely determined by f and x, and is called the *derivative* of f at x and denoted by $f'()$ or $Df(x)$. We write (1) as

$$f(x+h) = f(x) + Df(x)(h) + \phi(h) \tag{2'}$$

or more simply as

$$f(x+h) = f(x) + Df(x)(h) + o(\|h\|) . \tag{3}$$

If f is differentiable at each point $x \in U$, we say f is *differentiable* on U. In this case Df is a mapping:

$$Df : U \to L(E, F)$$

from U into the space of continuous linear maps $E \to F$; thus with each $x \in U$ we associate the linear mapping $Df(x) \in L(E, F)$. If Df is continuous, we say that f is *of class* C^1 or simply: f *is* C^1. We can inductively define f to be of class C^p if all derivatives $D^k f$ exist and are continuous for $0 \le k \le p$ understanding that $D^0 f$ means f itself.

Properties of the derivative.

(1) Suppose $U \subset E$ is an open set, $f, g : U \to F$ are mappings differentiable at $x \in U$. Then $f + g$ is differentiable at x and

$$D(f+g)(x) = Df(x) + Dg(x) .$$

If $c \in \mathbf{R}^1$, then $D(cf)(x) = cDf(x)$.

(2) **Chain Rule.** Suppose U and V are open sets in the Banach spaces E and F respectively, and f, g are mappings $\ni f : U \to V$, and $g : V \to G$, where G is also a Banach space. Let $x \in U$, and suppose f is differentiable at x, and g differentiable at $g(x)$. Then $g \circ f$ is differentiable at x and

$$D(g \circ f)(x) = Dg\big(f(x)\big) \circ Df(x) .$$

(3) Suppose $A : E \to F$ is a continuous linear map. Then A is differentiable at every point of E and $DA(x) = A(x)$ for every $x \in E$.

Also, if $U \subset E$ is an open set, $f : U \to F$ is differentiable, and $A : F \to G$ a continuous linear mapping, then $D(A \circ f)(x) = A \circ Df(x)\ \forall x \in U$, and for every $v \in E$, $D(A \circ f)(x)v = A\big(Df(x)v\big)$.

(4) If f is a differentiable mapping from a closed interval $[a, b]$ to F with zero derivative everywhere on $[a, b]$ then $f =$ constant on $[a, b]$.

Before stating further properties we shall explain a suitable theory of integration in one (real) variable for some of the formulae in differential calculus. The concept of integral here is arrived at as follows. Let $[a, b]$ be a closed interval, E a Banach space. Then a mapping $f : [a, b] \to E$ is called a *step map* if $\exists$ partition

$$\mathsf{P} : a = a_0 < a_1 < a_1 < \cdots < a_n \ = \ b$$

and elements $v_1, \ldots, v_n \in E \ni$ if t lies in the open interval (a_i, a_{i+1}) then $f(t) = v_{i+1}$, $i = 0, 1, \ldots, (n-1)$. We then say that f is step with respect to P. The concept of a refinement of a partition is the usual, and if f, g are two step maps of $[a, b]$ into E then $\exists$ partition $\mathsf{P} \ni$ both f and g are step w.r.t P. We see that the step maps form a subspace of the space of all bounded maps. We endow this space with the sup norm.

The integral of a step map f w.r.t to a partition P is defined by

$$I_{\mathsf{P}}(f) \ = \ \sum_{i=1}^{n} (a_i - a_{i-1})v_i$$

using the above notation. We see that this definition of the integral of a step map is independent of P, and we write simply $I(f)$ or $\int_a^b(f)$ to indicate the interval $[a, b]$. I is linear and $\|I(f)\| \ \le \ (b - a)\|f\|$. Hence I is continuous, with bound $b - a$. Then I is extended to the closure of the space of step maps by linear extension. If f lies in this

closure, we denote the integral $I(f)$ by $\int_a^b f$, and call it the integral and we then say that f is integrable. If $a \leq c \leq b$, then we verify that

$$\int_a^b f = \int_a^c f + \int_c^b f .$$

If $a \leq c < d \leq b$, then we define $\int_d^c f = -\int_c^d f$, and we verify that the formula $\int_c^e f = \int_e^d + \int_d^e f$ holds for any three points c, d, e in any order lying in an interval on which f is in the closure of the space of step maps.

A continuous map is uniformly continuous on a compact set, hence one concludes that the continuous maps of $[a, b]$ into E are in the closure of the space of step maps. Thus the integral is defined on continuous maps.

Suppose $E = E_1 \times E_2 \times \cdots \times E_n$ is a product of Banach spaces and a map $f : [a, b] \to E$ is an n-tuple $f = (f_1, \ldots, f_n)$ where $f_i : [a, b] \to E_i, i = 1, \ldots, n$, then

$$\int_a^b f = (\int_a^b f_1, \ldots, \int_a^b f_n) ,$$

if these integrals exist. If $E = \mathbb{R}^1$, and $f \geq 0$ then $\int_a^b f \geq 0$. Also if $f : [a, b] \to E$ is integrable, and $A : E \to F$ is a continuous linear map, then $\int_a^b A \circ f = A \circ \int_a^b f$.

Now going back to the differentiable calculus, the Fundamental theorem of the calculus holds just as in the real valued case as we find in the next result.

(5) If $f : [a, b] \to E$ is integrable and f is continuous at a point $c \in [a, b]$. Then the mapping $t \to \int_a^t f = \phi(t)$ is differentiable at c and its derivative at c is $f(c)$.

(6) If $I = [a, b]$, $\alpha(\cdot) : I \to L(E, F)$ is a continuous mapping and $y \in E$, then

$$\int_a^b \alpha(t) y dt = \int_a^b \alpha(t) dt \cdot y$$

where the right side means the application of the linear map $\int_a^b \alpha(t)dt$ to the vector $y \in E$.

(7) Let $U \subset E$ be open and $x \in U$. Suppose $f : U \to F$ is a C^1-map, and suppose the linear segment $\{x + ty, 0 \leq t \leq 1\}$ is contained in U. Then

$$f(x+y) - f(x) = \int_0^1 Df(x+ty) \cdot y dt = \int_0^1 Df(x+ty)dt \cdot y \ .$$

(8) Let $U \subset E$ be an open set and $x, z \in U \ni$ the line segment $I = \{(1-t)x + tz \mid 0 \leq t \leq 1\} \subset U$. If $f : U \to F$ is C^1 then

$$\|f(z) - f(x)\| \leq \|z - x\| \ \sup\{\|f'(v)\| \mid v \in I\} \ .$$

Higher order derivatives

Suppose $U \subset E$ is an open set and let $f : U \to F$ be differentiable in U. Then the derivative $Df(\cdot)$ is a mapping $Df(\cdot) : U \to L(E, F)$ the space of continuous linear maps from E to F (which is again a Banach space). Thus the second derivative is defined as the mapping which assigns to each $x \in U$, the derivative of Df, and if this derivative exists (denoted by $D^2 f(x)$), then

$$D^2 f(\cdot) : U \to L\big(E, L(E, F)\big) \ .$$

The right side here is identified with the space $L(E, E; F)$, the space of continuous bilinear maps $E \times E \to F$, and this space is for convenience denoted by $L^2(E; F)$.

(9) If $U \subset E$ is open and $f : U \to F$ is twice differentiable $\ni D^2 f(\cdot)$ is continuous in U, then $D^2 f(x)$ is symmetric for each $x \in U$.

And now we define derivatives of higher order than two by induction. If the $(p-1)^{th}$ derivative $D^{p-1}f(\cdot)$ exists in U then the p^{th} derivative is defined to be the mapping which assigns to each $x \in U$ the derivative of $D^{p-1}f$ at x and if this exists (denoted by $D^p f(x)$), then

$$D^p f(\cdot) : U \to L\big(E, L(E, \ldots)\big) .$$

The right side is identified with the space $L(\underbrace{E, \ldots, E}_{p}; F)$ and denoted by $L^p(E; F)$.

If $D^p f(x)$ exists for each $x \in U$ and if $D^p f(\cdot) : U \to L^p(E; F)$ is continuous for each $k = 0, 1, \ldots, p$, then we say f is C^p in U.

The p^{th} derivative D^p is linear, i.e.,

$$D^p(f+g) \ = \ D^p f + D^p g ,$$

$D^p(cf) \ = \ cD^p f$, c being a constant. Also if $p \ = \ q + r$, and if $D^p f(x)$ exists then $D^q D^r f(x) = D^p f(x) = D^r D^q f(x)$.

(10) If $U \subset E$ is open, and $f : U \to F$ is of class C^p, then for each $x \in U$, the mapping $D^p f(x)$ is multilinear symmetric.

(11) **Taylor's formula**. Suppose U is open in E and $f : U \to F$ is of class C^p. Let $x \in U$, and $y \in E \ni$ the segment $\{x + ty \mid 0 \le t \le 1\}$ is contained in U. Then

$$f(x+y) = f(x) + Df(x) \cdot y + \frac{1}{2!} D^2 f(x) y^{(2)} + \cdots + \frac{1}{(p-1)!} D^{p-1} f(x) \cdot y^{(p-1)} + R_p$$

where the remainder R_p is given by

$$R_p \ = \ \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} D^p f(x+ty) \cdot y^{(p)} dt .$$

We shall conclude this appendix with a result on differentiation of sequences of mappings.

(12) Suppose $U \subset E$ is open, and $\{f_n\}$ is a sequence of C^1-mappings $U \to F$. Suppose $\{f_n\}$ converges pointwise to f, and also the sequence $\{Df_n\}$ converges uniformly to a mapping $g : U \to L(E, F)$. Then Df exists and $= g$.

This Page Intentionally Left Blank

APPENDIX 3

Differentiable Banach manifolds

We shall collect in this appendix the basic ideas, definitions and results concerning differentiable Banach manifolds which are needed in this book. For more details and proofs the reader should refer to Lang [32], or Palais [47].

Definition. *Let X be a topological space. A chart in X is a homeomorphism ϕ defined on an open set $D(\phi) = U \subset X$ onto an open set in a Banach space V (or onto an open set in a half-space of V).*

If ϕ, ψ are charts in X defined on open sets $D(\phi) = U_1$, $D(\psi) = U_2$ respectively and $U = U_1 \cap U_2$, then ϕ, ψ are *C^k-related* if $\psi \circ \phi^{-1}$ is a C^k-isomorphism of $\phi(U)$ onto $\psi(U)$.

A *C^k-atlas* for X is a collection $\mathcal{A}$ of charts in X which are pairwise C^k-related such that $X = \bigcup_{\phi \in \mathcal{A}} D(\phi)$. A *complete C^k-atlas* is one which is maximal in the ordering by inclusion.

We shall need the following simple facts which we shall state as Lemmas.

Lemma. *Let $\mathcal{A}$ be an atlas for X and ϕ, ψ charts in X each of which as C^k-related to each chart in $\mathcal{A}$. Then ϕ, ψ are C^k-related.*

Lemma. *Any C^k-atlas $\mathcal{A}$ for X is included in a unique complete C^k-atlas $\tilde{\mathcal{A}}$ viz. $\tilde{\mathcal{A}} = \{\phi$ is a chart in X C^k-related to each chart of $\mathcal{A}\}$.*

$\tilde{\mathcal{A}}$ in the preceding Lemma will be refered to as the C^k-completion of the atlas $\mathcal{A}$.

Definition. *A C^k-manifold M is a pair $(X, \mathcal{A})$ where X is a Hausdorff space and $\mathcal{A}$ is a complete C^k-atlas for X; the underlying topological space X will often be denoted by M. If $p \in M$, then an element $\phi \in \mathcal{A}$ with $p \in D(\phi)$ will be called a chart for M at p.*

On the other hand if $\mathcal{A}$ is any C^k-atlas for X then for any integer $m \leq k$, by the *C^m-manifold defined by $\mathcal{A}$* we shall mean the manifold $(X, \tilde{\mathcal{A}}^{(m)})$, $\tilde{\mathcal{A}}^{(m)}$ being the C^m-completion of $\mathcal{A}$. If M is a C^k-manifold $(k \geq 1)$ the *boundary* of M is denoted by ∂M, shall mean the set

$$\{p \in M \,\Big|\, \exists \text{ chart } \phi \text{ for } M \text{ at } p \text{ mapping } D(\phi) \quad \text{onto an open}$$
$$\text{set of a half-space} \quad H \subset V, \text{ with } \phi(p) \in \partial H\} .$$

This leads to the next lemmas.

Lemma. *If M is a C^k-manifold $(k \geq 1)$ and $p \in \partial M$, and ϕ is any chart for M at p, then ϕ maps $D(\phi)$ onto an open set of a half-space H, with $\phi(p) \in \partial H$.*

Lemma. *If $M = (X, \mathcal{A})$ is a C^k-manifold $(k \geq 1)$ then the set $\{\psi \,\big|\, \psi = \phi\big|_{\partial M}, \phi \in \mathcal{A}\}$ is a C^k-atlas for ∂M.*

Definition. *If M, N are C^k-manifolds, and $f : M \to N$ is a mapping, then f is C^k near p if $\exists$ charts ϕ for M at p, ψ for N at $f(p) \ni \psi \circ f \circ \phi^{-1}$ is C^k at $\phi(p)$; and f is C^k on M if f is C^k near every point $p \in M$.*

This definition then yields the following theorem.

Theorem. *Suppose M, N are C^k-manifolds, and $f : M \to N$ is a mapping. Then f is C^k on M if and only if $\psi \circ f \circ \phi^{-1}$ is C^k for every choice of charts ϕ for M, ψ for N.*

Corollary. *Suppose V is a Banach space, H is a half-space; then the identity mappings I_V and I_H are C^k-atlases for V for V and H respectively, and thus V, H are C^k-manifolds in the preceding sense, for any $k \geq 1$.*

Definition. *Suppose M is a C^k-manifold and N is a subspace of M. A chart ϕ for M*

will be said to admit restriction to N *if* $\exists$ *closed linear subspace* $W \subset V \ni \phi\Big|_N$ *is a chart mapping* $D(\phi) \cap N$ *homeomorphically onto an open subset of* W *or a half-space of* W.

Lemma. *Suppose* M *is a* C^k*-manifold,* N *a subspace of* M*, and* ϕ, ψ *are charts for* M *which admit restrictions to* N*, then* $\phi\Big|_N, \psi\Big|_N$ *are* C^k*-related.*

Lemma. *If* N *is a* C^k*-submanifold of the* C^k*-manifold* M*, and* $\mathcal{B} = \{$*charts* ϕ *for* $M \Big| \phi$ *admits a restriction to* $N\}$*, then* $\tilde{\mathcal{B}} = \{\phi\Big|_N \Big| \phi \in \mathcal{B}\}$ *is a* C^k*-atlas for* N*; in such a case we shall denote the manifold obtained by the* C^k*-completion of the atlas by* N*, also;* N *is called a properly imbedded submanifold.*

Lemma. *Suppose* N *is a* C^k*-submanifold of the* C^k*-manifold* M*. Then the inclusion map* $i_M : N \to M$ *is a* C^k*-isomorphism. If further,* $f : M \to M'$ *is a* C^k *map then so is* $f\Big|_N = f \circ I_N$*. If* $f : M \to M'$ *is a* C^k*-isomorphism and* $\mathcal{O}$ *is open in* M*, then* $\mathcal{O}$ *is a* C^k*-submanifold of* M*.* ∂M *is always a* C^k*-submanifold of* M*.*

Lemma. *Suppose* M, N *are submanifolds, one of them without a boundary. If* ϕ *(respectively* ψ*) is a chart for* M *(respectively* N*), then* $\phi \times \psi : D(\phi) \times D(\psi) \to$ *(target space of* ϕ*)* $\oplus$ *(target space of* ψ*) is a chart in* $M \times N$*. The set of such* $\phi \times \psi$ *is a* C^k*-atlas for* $M \times N$ *and the* C^k*-manifold this atlas defines will be denoted also by* $M \times N$*.*

Tangent spaces.

Suppose M is a C^k-manifold ($k \geq 1$), and $\mathcal{A}_p$ the set of charts at $p \in M$. We shall consider triples (U, ϕ, a), where $(U, \phi) \in \mathcal{A}_p$ and a is an element of the vector space in which ϕU lies. Two such triples (U, ϕ, a), (V, ψ, b) are said to be *equivalent* $(\sim)$ if $D(\psi \circ \phi^{-1})(\phi(p))a = b$. This agreement defined an equivalence relation $\sim$ (by the chain rule).

A *tangent vector at* p means an equivalence class of such triples under the equivalence relation $\sim$, and the set of all such tangent vectors at p is called the *tangent space of* M at p and denoted by T_pM. Each chart (U, ϕ) at p induces a Banach space structure on T_pM, which is independent of the chart selected.

Suppose U, V are open sets in Banach spaces, then to every mapping $f : U \to V$ which is say C^k-smooth ($k \geq 1$), we associate its derivative $Df(x)$. If $f : X \to Y$ is a C^p-smooth mapping of a manifold X into another manifold Y, then by means of charts we can meaningfully define the derivative of f on each chart at p as a mapping

$$T_pf : T_p(X) \to T_{f(p)}(Y) \; .$$

This map is the unique linear mapping with the property: suppose (U, ϕ) is a chart at p and (V, ψ) is a chart at $f(p) \ni f(U) \subset V$ and $\bar{a}$ is a tangent vector at p represented by a in the chart (U, ϕ), then $T_pf(\bar{a})$ is the tangent vector at $f(p)$ represented by $Df_{U,V}(p)a$. This map T_pf is linear and continuous for the Banach space structure placed on $T_p(X)$ and $T_{f(p)}(Y)$.

For convenience we shall sometimes write f_{*p} instead of T_pf. This mapping T has the properties: if $f : X \to Y$, and $g : Y \to Z$ are C^p-mappings ($p \geq 1$) then

$$T_p(f \circ g) \; = \; T_{f(p)}(g) \circ T_p(f) \; ,$$

and

$$T_p(id) \; = \; id \; .$$

Two major concepts to be noted are the concepts of *immersion* and *submersion*:

Suppose X, Y are manifolds modelled on Banach spaces, and suppose $f : X \to Y$ is a mapping. Let $p \in X$.

(a) f is an *immersion at* p if $\exists$ open neighbourhood X_1 of p in $X \ni$ the restriction of f to X_1 induces a homeomorphism of X_1 onto a submanifold of Y. f is an *immersion* if it is an immersion at every point.

(b) f is a *submersion at* p if $\exists$ chart (U, ϕ) and a chart (V, ψ) at $f(p) \ni \phi$ gives a homeomorphism of U on a product $U_1 \times U_2$ where U_1, U_2 are open sets in some Banach spaces and $\ni$ the mapping

$$\psi f \phi^{-1} = f_{V,U} : U_1 \times U_2 \to V$$

is a projection. We see that the image of a submersion is an open subset.

The following is a useful criterion for immersions and submersions in terms of the derivative.

Theorem. *Suppose X, Y are C^p-manifolds ($p \geq 1$) modelled on Banach spaces, and suppose $f : X \to Y$ is a C^p-mapping. Let $x \in X$. Then*

(a) f is an immersion at x if and only if $T_x f$ is one-one and its image has a complement;

(b) f is a submersion at x if and only if $T_x f$ is onto and its kernel has a complement.

The concept of a mapping which is transversal over a submanifold needs to be clarified. A mapping $f : X \to Y$ is said to be *transversal* over the submanifold $W \subset Y$ is the following condition is satisfied:

Let $x \in X \ni f(x) \in W$. Let (V, ψ) be a chart at $f(x) \ni \psi : V \to V_1 \times V_2$ is a homeomorphism onto a product with $\psi\big(f(x)\big) = (0, 0)$ and $\psi(W \cap V) = V_1 \times 0$. Then $\exists$

open neighbourhood U of $x \ni$ the composite mapping

$$(pr \circ \psi \circ f) : U \to V_2$$

is a submersion.

In particular if f is transversal over W then $f^{-1}(W)$ is a submanifold of X, since the inverse image of 0 by the local composite $(pr \circ \psi \circ f)$ is equal to the inverse image of $W \cap V$ by ψ.

The following is a characterisation of transversal maps in terms of tangent spaces.

Theorem. *Suppose X, Y are C^p-manifolds ($p \geq 1$) modelled on Banach spaces, and suppose $f : X \to Y$ is a C^p-mapping and W a submanifold of Y. The mapping f is transversal over W if and only if for each $x \in X \ni w = f(x) \in W$ the composite mapping*

$$T_x(X) \xrightarrow{T_x f} T_W(Y) \to T_W(Y)/Y_w(W)$$

is onto and its kernel has a composite.

APPENDIX 4

Probability theory

It has been recognised that a satisfactory and rigorous presentation of probability theory can be given only in the setting of modern measure theory and abstract integration. Several excellent accounts of measure theory are now available, starting with S. Saks [53], Halmos [16], Loève [37] and more recently, in the monograph of Wheeden and Zygmund [65]. We shall assume familiarity with basic measure theory and integration. In this appendix we shall review the basic concepts and results in probability theory which are needed in this monograph. For very elementary explanations of simple probabilistic distributions such as e.g., the binomial distribution arising in connection with a succession of coin-tossing experiments, we refer the reader to e.g., M. Rosenblatt [51].

The basic setting for probabilistic results is a *probability space*, i.e., a triple $(\Omega, \mathcal{B},)$ where Ω is a space of points, $\mathcal{B}$ is a Borel field (sometimes the word σ-field is also used) of special subsets of Ω called *measurable* sets, and P is a probability measure on $\mathcal{B}$, i.e., $P(\Omega) = 1$. The subsets of $\mathcal{B}$ are the "random events" on which a "probability" viz. P, is defined. A *random variable* (r.v.) is a function $X(\omega)$ measurable w.r.t. $\mathcal{B}$; it is thought of as a possible observable in an experiment whose outcome is governed by the measure P. The integral

$$E(X) = \int X(\omega)P(d\omega)$$

(if it exists) is called the *mean or expectation of* X.

Suppose X is a random variable with a finite absolute moment: $E(|X|) < \infty$. A

very useful inequality is *Chebyshev's inequality*: for $c > 0$,

$$P(|X| \geq c) \leq \frac{1}{c} E(|X|) .$$

It is useful in obtaining bounds on the amount of probability or probability mass in the "tail" of distribution. A variant of Chebyshev's inequality is: if $\alpha > 0, c > 0$, then

$$P(|X| \geq c) \ \leq \frac{1}{c^\alpha} E|X|^\alpha$$

provided the expectation on the right side exists.

Jensen's inequality is also a useful inequality. Suppose ϕ is a real-valued continuous convex function of a real variable. Convex means for every pair of points $x, x \in \mathbf{R}^1$,

$$\phi\left(\frac{x + x'}{2}\right) \ \leq \ \frac{1}{2}\phi(x) + \frac{1}{2}\phi(x') .$$

If ϕ is continuous then ϕ is convex if and only if for each $x_0 \ \in \mathbf{R}^1$, there is a number $\lambda(x_0)$ such that for all $x \in \mathbf{R}^1$,

$$\lambda(x_0)(x - x_0) \ \leq \phi(x) - \phi(x_0) .$$

The inequality of Jensen which we are referring to here, asserts that

$$E\phi(X) \ \geq \ \phi(EX) .$$

Events (measurable sets) $A_1, \ldots, A_n$ are *independent* if the following is true: writing $A_i^{(0)} - A_i, A_i^{(1)} = \mathcal{C}(A_i)$ (i.e., the complement of A_i), we have

$$P\left(\bigcap_{i=1}^{n} A_i^{(m_i)}\right) \ = \ \prod_{i=1}^{n} P(A_i^{(m_i)})$$

where $m_i = 0, 1$, for $i = 1, \ldots, n$. An arbitrary family of events $(A_\alpha)_{\alpha \in I}$ is *independent* if every finite subset is independent.

The *Borel-Cantelli Lemma* gives us the probability of the set $\{A_i \ i.o.\}$ of points lying in an infinite number of sets $A_j, j = 1, 2, \ldots$. The set $\{A_i \ i.o.\}$ is defined to be the set $\bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j$, or the *superior limit* of A_n denoted by $\limsup \ A_n$. This lemma then, is as follows.

Borel-Cantelli Lemma. *Let $P(A_i)$ be the probability of the event A_i, $i = 1, 2, \ldots$. If $\sum_{i=1}^{\infty} P(A_i) < \infty$ then $P(A_i \ i.o.) = 0$. If the events $A_i, i = 1, 2, \ldots$ are independent, then $\sum_{i=1}^{\infty} P(A_i) = \infty$ implies that $P(A_i \ i.o.) = 1$.*

There is a basic zero-one law also due to Kolmogorov, somewhat similar to the Borel-Cantelli Lemma. Some explanation is necessary before this law can be stated. The terminology "A_n's (occurring) infinitely often" corresponds to the fact that $\limsup A_n$ is the set of all those events which belong to infinitely many A_n, or equivalently, to some of "the $A_n, A_{n+1}, \ldots$, however large be n", i.e. the so-called "tail" of the sequence $\{A_n\}$. To the "tail" of the sequence $\{A_n\}$ corresponds the "tail" of the sequence $\{\psi_{A_n}\}$ of their characteristic functions.

More generally the "tail" of a sequence of r.v.'s is explained as follows. Let $\{X_n\}$ be a sequence of r.v.s, and let $\mathcal{B}(X_n), \mathcal{B}(X_n, X_{n+1}), \ldots, \mathcal{B}(X_n, X_{n+1}, \ldots)$, $\mathcal{B}(X_{n+1}, X_{n+2}, \ldots,)$ be the sub-σ-fields of events induced by the random functions within the brackets. The concept of limsup $\mathcal{B}(X_n)$ is defined precisely as follows. The sequence $\mathcal{B}(X_n), \mathcal{B}(X_n, X_{n+1}) \ldots$ is a non-decreasing sequence of σ-fields, the minimal σ-field over its supremum or union, is $\mathcal{B}(X_n, X_{n+1}, \ldots)$, which is also loosely written as

"$\sup_{m\geq n} \mathcal{B}(X_m)$". Furthermore, the sequence $\mathcal{B}(X_n, X_{n+1}, \ldots)$, $\mathcal{B}(X_{n+1}, X_{n+2}, \ldots)$, ... is a non-increasing sequence of σ-fields, its limit or intersection is a σ-field contained in $\mathcal{B}(X_n, X_{n+1}, \ldots)$ however large n may be, and loosely denoted by "limsup $\mathcal{B}(X_n)$". This σ-field $\mathcal{C}$ is called the "tail σ-field of the sequence $\{X_n\}$", or "the sub σ-field of events induced by the tail of the sequence $\{X_n\}$", and its elements are called "tail events", and functions (finite or not) which are $\mathcal{C}$-measurable are called "tail functions". Thus limsup X_n, as also the limits inferior and superior of the sequence $\{\frac{1}{n}\sum_{i=1}^{n} X_i\}$ are tail functions, while the sets of convergence of these sequences, or the set of convergence of ΣX_n, are tail events. Kolmogorov's law is the following.

Kolmogorov's Zero-one Law. On a sequence of independent r.v.'s, the tail events have probability either 0 or 1 and the tail functions are degenerate.

In our considerations in this monograph the concept of a stochastic process in an important one. Very simply stated, a stochastic process is a mathematical model of a process which occurs in nature. In other words, in non mathematical terms we can describe a stochastic process as a process evolving in time and subject to probabilistic laws. If we make numerical observations as the process continues in time, these observations give some idea of the evolution of the process. Thus it is customary to define a *stohcastic process* as a family of r.v.'s $\{X_t^{(\omega)} \mid t \in \tau\}$ (cf. [51]). Here τ is the "time" range and $X_t(\omega)$ in practice is the observation at time $t \in \tau$; the "ω" serves to denote the dependence of the t^{th} observation (or "observable") on random causes i.e. which are subject to a probabilistic law.

In practice, if a concrete process in the natural world is being observed, we can only

observe it at a finite set of instants of time, one group of observations being made at instants $t_1, t_2, \ldots, t_n$, say, another group of observations perhaps at instants $t'_1, t'_2, \ldots, t'_m$; and so on. We know the space in which the values of these observations lie. Also we may have good reason to conclude that each finite group of "random" observations, say at $\tau_n = (t_1, t_2, \ldots, t_n)$, is subject to a definite probabilistic law, associated with a probability measure μ_{τ_n}. And then we would like to conclude that all these observables observed at different instants of time in varying (finite) groups are actually random variables (as defined earlier) on one probability space in which these finite dimensional measures μ_{τ_n} can be realised. A basic theorem of Kolmogorov describes a condition under which this does really occur. This condition is called the "consistency condition".

Before stating Kolmogorov's theorem we shall be more explicit about the space S in which our observables take values. S is assumed to be a σ-compact Hausdorff space with the Borel field $\mathcal{A}$ generated by the topology on S. τ is the time interval over which the observables are observed. At each instant of time $t \in \tau$, S_t is defined to be $= S$, and $\mathcal{A}_t$ is taken to be the Borel field $\mathcal{A}$. If $\tau_n = (t_1, \ldots, t_n)$ is a finite subset of τ, then $\mathcal{B}_{\tau_n}$ is the product Borel field $\prod_{t_i \in \tau_n} \mathcal{A}_{t_i}$.

Kolmogorov's Extension Theorem. Suppose the range space of an observable to be a σ-compact Hausdorff space S, with the Borel field $\mathcal{A}$ generated by the topology on S. Assume that for each finite subset τ_n of τ there is given a probability measure μ_{τ_n} on $\mathcal{B}_{\tau_n}$ regular with respect to the product topology on $\prod_{t_i \in \tau_n} S_{t_i}$. Suppose the family $\{\mu_{\tau_n}\}$ of measures satisfies Kolmogorov's consistency condition:

Consistency condition: every pair of measures $\mu_{\tau_n}, \mu_{\tau_m}$ corresponding to finite sub-

sets τ_n, τ_m of τ, must agree on the Borel field $\mathcal{B}_{\tau_n \cap \tau_m}$. (**Note:** here, denoting as usual by $\emptyset$ the empty set, $\mathcal{B}_\emptyset$ is understood to be the trivial Borel field consistency of only the null set and the whole space.)

Contention: Then there exists a certain probability space and there exists a stochastic process $\{X_t(\omega), t \in \tau\}$ on this probability space realising these finite dimensional measures $\mu_{\tau_n}, \tau_n \subset \tau$.

A little more explanation regarding this extension theorem of Kolmogorov appears in order. Let $S^\tau = \prod_{t \in \tau} S_t$, the product space of points $\omega = (\omega_t, t \in \tau)$ whose t^{th} co-ordinate is $\omega_t \in S_t$. We identify a finite-diemnsional set $B \in B_{\tau_n}$ on $S^{\tau_n} = \prod_{t \in \tau_n} S_t$ with the set $B \times \prod_{t \in \tau - \tau_n} S_t = B \times S^{\tau - \tau_n}$, and set

$$\mu(B \times S^{\tau-\tau_n}) = \mu_{\tau_n}(B), \qquad B \in \mathcal{B}_{\tau_n} .$$

Thus $\mathcal{B}_{\tau_n}$ is identified with the Borel field of "cylinder sets" $B \times S^{\tau-\tau_n}$, $B \in \mathcal{B}_{\tau_n}$ on S^τ which we shall still call $\mathcal{B}_{\tau_n}$. Then using the consistency restriction on the measures $\{\mu_{\tau_n}\}$, the definition of μ on all sets $B \times S^{\tau-\tau_n}$, $B \in \mathcal{B}_{\tau_n}$ for all finite sets τ_n of τ, is possible because of the consistency condition on the set $\{\mu_{\tau_n}\}$. Then μ is a non-negative additive set function, with $\mu(S^\tau) = 1$, on the field $\mathcal{F} = \bigcup_{\tau_n \subset \tau} \mathcal{B}_{\tau_n}$ obtained by taking the union of the Borel fields $\mathcal{B}_{\tau_n}$ for all finite subsets τ_n of τ.

Next we invoke the following extension result due to Caratheodory.

Caratheodory's extension theorem: Let ν be a non-negative additive set function on a field $\mathcal{F}$ of sets of a space Λ, with $\nu(\Lambda) = 1$. If ν is continuous at the null set $\emptyset$ then $\exists|$ extension $\bar{\nu}$ of $\nu(\bar{\nu}(B) = \nu(B) \forall B \in \mathcal{F})$, on the Borel field $\mathcal{B}$ generated by the field $\mathcal{F}$.

Using this extension theorem we finally conclude that the additive measure μ on $\mathcal{F}$ can be extended to the Borel field $\mathcal{B}$ generated by $\mathcal{F}$, as a probability measure. The required stochastic process is obtained by defining

$$X_t(\omega) = \omega_t, \qquad t \in \tau$$

as the random variables of the stochastic process.

This Page Intentionally Left Blank

BIBLIOGRAPHY

[1] Abraham, R. and Robbin, J., *Transversal mappings and flows*, (1967), W.A. Benjamin, New York.

[2] Aron, R. and Prolla, J.B., Polynomial approximation of differentiable functions on Banach spaces, *J. fur Reine Angervendte Math.*, **313** (1980), 195-210.

[3] Balazs, Katherine, Approximation by Bernstein-type rational functions on the real line, *Acta Math. Hung.*, **46** (1-2) (1985), 195-204.

[4] Bernstein, S.N., Démonstration du theorème de Weierstrass fondée sur le calcul de probabilité, Proc. Kharkov Math. Soc. XIII, 1912.

[5] Bonic, R. and Frampton, J., Smooth functions on Banach manifolds, *J. Math. Mech.*, **15** (1966), 877-895.

[6] Butzer, P.L. and Behrens, H., *Semi-groups of operators and approximation*, (1967), Springer-Verlag, New York.

[7] Cisielski, Z., Hölder conditions for realisations of Gaussian processes, Amer. Math. Soc., Transactions, Vol. 99 (1961), 403-413.

[8] Desolneux-Moulis, Nicole, Approximation de fonctions différentiables sur certains espaces de Banach, Ann. Inst. Fourier, Grenoble, 21, 4 (1971), 293-345.

[9] Dieudonné, J., *Foundations of modern analysis*, (1960), Academic Press, New York.

[10] Eells, J. and McAlpin, J., An approximate Morse-Sard theorem, *J. Math. Mech.*, **17** (1967), 1055-1064.

[11] Enflo, P., Banach spaces which can be given a uniformly convex norm, *Israel J. Math.*, **13** (1972), 281-288.

[12] Federer, H., *Geometric measure theory*, (1969), Springer-Verlag, New York.

[13] Feller, W., *An introduction to probability theory and its applications, Vol II*, (1957), John Wiley, New York.

[14] Gnendenko, B.V. and Kolmogorov, A.N., *Limit distributions for sums of independent random variables*, (1954), Addison Wesley Publishing Co., Cambridge, Mass.

[15] Gohberg, I.C. and Krein, M.G., Basic properties of defect numbers and indices of linear operators, *Uspehi Mat. Nauk.* 12, 2 (74) (1957), 43-118.

[16] Halmos, P., *Measure theory*, (1950), Van Nostrand, New York.

[17] Heble, M.P., Approximation of differentiable functions on a Hilbert space, *C.R. Math. Rep. Acad. Sci. (Can.) V* (1983), 179-183.

[18] Heble, M.P., Approximation of differentiable functions on a Hilbert space II, Contemporary Mathematics, Amer. Math. Soc., Providence, R.I., Vol. 54 (1986), 17-33.

[19] Heble, M.P., Approximation of differentiable functions on a Hilbert space III, Math. Ann., Vol. 282 (1988), 473-484.

[20] Hille, E. and Phillips, R., Functional analysis and semi-groups, (Amer. Math. Soc. Colloq. Publ. Vol. 31, revised 1974 c. 1957, Providence, R.I.).

[21] Hirsch, M., *Differential topology*, (1976), Springer-Verlag, New York.

[22] Ito, K. and McKean, H.P., *Diffusion processes and their sample paths*, Springer-Verlag, Berlin, Heidelberg and New York.

[23] Ito, K. and Nisio, M., On stationary solutions of stochastic differential equations, *J. Math. Kyoto Univ.*, **4** (1964), 1-75.

[24] James, R.C., Some self-dual properties of normed linear spaces, Symp. Infinite dimensional topology, Ann. Math. Studies, Vol. 69 (1972), 159-175.

[25] Johnson, W.B., Rosenthal, H.P. and Zippin, M., On bases, finite-dimensional decompositions and weaker structures in Banach space, *Israel J. Math.*, **9** (1971), 488-506.

[26] Kahane, J.-P., Séries de Fourier aléatoires, Séminaire Bourbaki, 12ᵉ année, 1959/60, no. 200.

[27] Kallianpar, G., *Stochastic filtering theory* , (1980), Springer-Verlag, New York.

[28] Khan, R.A., Some probabilistic methods in the theory of approximation operators, Acta Math. Acad. Sci. Hung., Tomus 35 (1-2), (1980), 193-203.

[29] Köthe, G., *Topological vector spaces I*, (1969), Springer-Verlag, New York.

[30] Kurzweil, J., On approximation in real Banach spaces, *Studia Math.*, **14** (1954), 214-231.

[31] Kurzweil, J., On approximation in real Banach spaces by analytic operations, *Studia Math.*, **17** (1957), 124-129.

[32] Lang, S., *Differential manifolds*, (1972), Addison Wesley Publ. Co., Reading, Mass.

[33] Lang, S., *Real Analysis*, (1969), Addison Wesley Publ. Co., Reading, Mass.

[34] Lesmes, J., On the approximation of continuously differentiable functions in Hilbert spaces, *Rev. Coll. Mat.* **8** (1974), 217-233.

[35] Levy, P., *Processus stochastiques et mouvement brownien*, (1948), Gauthier-Villar, Paris.

[36] Llavona, J.G., *Notas de Matematica 130*, (1986), North Holland Publ. Co.

[37] Loève, M., *Probability theory*, (1960), Van Nostrand, New York.

[38] Lorentz, G.G., *Bernstein polynomials*, (1953), Univ. of Toronto Press, Toronto.

[39] Lusternik, R.A. and Sobolev, V.J., *Elements of functional analysis*, (1986 c. 1961), Gordon and Breach, New York.

[40] Malgrange, B., *Ideals of differentiable functions*, (1964), T.I.F.R. Studies in Mathematics; Oxford University Press, Bombay.

[41] Mergleyan, S.N., Weighted approximation by polynomials, Amer. Math. Soc. Translations, Series 2, Vol. 10, (1958), 50-106.

[42] Munkres, J.R., Elementary differential topology, Annals of Mathematics Studies, Vol. 54, Princeton 1966.

[43] Nachbin, L., *Elements of approximation theory*, (1976), Robert E. Krieger Publ. Co., Huntington, New York.

[44] Nachbin, L., Sur les algèbres denses de fonctions differentiables sur une variété, *C.R. Acad. Sci.*, Paris, **288** (1949), 1549-1551.

[45] Nemirovski, A.M. and Semenov, S.M., On polynomial approximation of functions in a Hilbert space, Sbornik U.S.S.R. **21** (1973), 255-277.

[46] Orlicz, W., Berträge zur theorie der Orthogonal entwicklungen II, *Studia Math.*, **1** (1929), 241-255.

[47] Palais, R.S., Lectures on the differential topology of infinite-dimensional manifolds, Brandeis Univ., 1964-1965.

[48] Pollard, D., *Convergence of stochastic processes*, (1984), Springer-Verlag, New York.

[49] Restrepo, G., An infinite-dimensional version of a theorem of Bernstein, Proc. Amer. Math. Soc. 23 (1969), 193-198.

[50] Rosenblatt, M., *Markov processes, structure and asymptotic properties*, (1971), Springer-Verlag, New York.

[51] Rosenblatt, M., *Random processes*, Springer-Verlag, New York.

[52] Rudin, W., *Functional analysis*, (1973), McGraw Hill, New York.

[53] Saks, S., *Theory of the integral*, (1937), Hafner Publ. Co., New York.

[54] Schwartz, L., La fonction aléatoires du mouvement brownien, Séminaire Bourbaki, Fev. 1958, no. 161.

[55] Shiota, M., Approximation theorems for Nash mappings and Nash manifolds, Trans. Amer. Math. Soc., Vol. 293, no. 1, Jan. 1986, 319-337.

[56] Skorokhod, A.V., Limit theorems for stochastic processes, I, Theory of Probability and its Applications, Vol. I, no. 3, 1956, 261-290.

[57] Skorokhod, A.V., Limit theorems for stochastic processes with independent increments II, Theory of Probability and its Applications, Vol. II, no. 2, 1957, 138-171.

[58] Smale, S., An infinite-dimensional version of Sard's theorem, *Amer. J. Math.*, **87** (1965), 861-866.

[59] Steinhaus, H., Uber die Wahrscheinlichkeit dafur, dass der Konvergenzkreis einer Potenzreihe ihre naturliche Grenze ist, Math. Z., t. 31 (1930), 408-416.

[60] Stone, M.H., The generalized Weierstrass theorem, *Mat. Mag.*, **21** (1948), 167-184, 237-254.

[61] Sundaresan, K., Geometry and non-linear analysis in Banach spaces, *Pac. J. Math.*, **102** , no. 1 (1982).

[62] Sundaresan, K. and Swaminathan, S., *Geometry and non-linear analysis in Banach spaces*, Lect. Notes Math., Vol. 1131, (1986), Springer-Verlag, New York.

[63] Wells, J., Differentiable functions on c_0, *Bull, Amer. Math. Soc.*, **75** (1969), 117-118.

[64] Wittman, R., Approximations - und Fursetzungssätze für konvexe funktionen, *Arch. Math.* , **45** (1985), 463-467.

[65] Wheeden, R.L. and Zygmund, A., *Introduction to real analysis*, (1977), Marcel Dekker, New York.

[66] Whitney, H., Analytic extensions of differentiable functions defined in closed sets, *Trans. Amer. Math. Soc.*, **36** (1934), 63-89.

[67] Zygmund, A., *Trigonometric Series*, 2nd ed., Cambridge, University Press, 1959.

This Page Intentionally Left Blank

INDEX

This Page Intentionally Left Blank